DOCUMENT RESUME

ED 147 311 TM 005 705

AUTHOR Elford, George; And Others
TITLE Handbook of Management Standards for Awarding High School Diplomas.
INSTITUTION Massachusetts Advisory Council on Education, Boston.
PUB DATE Sep 75
NOTE 116p.

EDRS PRICE MF-$0.83 HC-$6.01 Plus Postage.
DESCRIPTORS Alternative Schools; Basic Skills; *Educational Alternatives; Evaluation Criteria; *Evaluation Methods; External Degree Programs; *Graduation Requirements; *Guides; Performance Based Education; Performance Tests; Program Evaluation; Public Schools; Secondary Education; *Senior High Schools; *Standards; Student Evaluation; Test Construction; Work Experience
IDENTIFIERS Alternatives to Standardized Testing; Experiential Learning; *Massachusetts

ABSTRACT

This handbook is designed to assist Massachusetts schools in utilizing members of the community, teachers, and administrators in developing a system of evaluation to monitor student progress and measure program effectiveness. Emphasis is placed on the establishment of quality control procedures for alternative learning experiences, such as community-based programs, action-learning, or experiential learning, and the application of alternative methods of evaluating such experiences. A description of graduation requirements in Massachusetts and current practices and needs in experiential learning provides the background for a discussion of setting management standards and minimal acceptable levels of performance for these programs. Five assessment methods that can be used either separately or in combination to assess student competencies for awarding high school credit are discussed. They are: testimony, committee review, performance assessment, product assessment, and objective and essay examinations. Descriptions of several nontraditional learning programs, and sample documents for a performance test are appended. (EVH)

****
* Documents acquired by ERIC include many informal unpublished *
* materials not available from other sources. ERIC makes every effort *
* to obtain the best copy available. Nevertheless, items of marginal *
* reproducibility are often encountered and this affects the quality *
* of the microfiche and hardcopy reproductions ERIC makes available *
* via the ERIC Document Reproduction Service (EDRS). EDRS is not *
* responsible for the quality of the original document. Reproductions *
* supplied by EDRS are the best that can be made from the original. *
****

ED147311

handbook of management standards

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Robert Kates

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC) AND USERS OF THE ERIC SYSTEM."

TM00

for awarding high school diplomas

mece a study conducted for the massachusetts advisory council on education

Cover design by Joseph L. Belica

HANDBOOK OF MANAGEMENT STANDARDS

FOR AWARDING HIGH SCHOOL DIPLOMAS

Study Directors:

Robert Kates
College Entrance Examination Board

John Fremer
Educational Testing Service

Handbook prepared by:

George Elford
John Fremer
Robert Kates
Kathleen Rice

a study conducted for the
Massachusetts Advisory Council
on Education
182 Tremont Street
Boston, Massachusetts 02111

September 1975

The Massachusetts Advisory Council on Education is an independent state agency created by special legislation (General Laws, Chapter 15, Section 1H) for the purpose of recommending policies designed to improve the performance of all public education systems in the Commonwealth. As such, the Advisory Council provides support for studies that will recommend policies promoting and facilitating the coordination, effectiveness and efficiency of these educational systems.

MASSACHUSETTS ADVISORY COUNCIL ON EDUCATION

Members of the Council
Mary Warner, Chairman; Engineer, Sunderland
Benjamin D. Fleet, Vice Chairman; President, Fleet Tire Company, Falmouth
Alford Dyson, Textile Workers Union of America, Westport
Oliver W. Kerr, Accounting Manager, New England Telephone Company, Springfield
Elaine Kistiakowsky, League of Women Voters of Massachusetts, Cambridge
Milton Paisner, General Manager, Electronic Products, Inc., Newburyport
Nina E. Scarito, M.D., Obstetrician, Methuen
Verne W. Vance, Jr., Attorney, Foley, Hoag and Eliot, Boston
Gregory R. Anrig, Commissioner of Education, ex officio
Roy Keith, Interim Chancellor, Board of Higher Education, ex officio

Legislative Consulting Committee
Walter J. Boverini, State Senator, Lynn
Nicholas J. Buglione, State Representative, Methuen
Edward L. Burke, State Senator, Framingham
Michael J. Daly, State Representative, Brighton
Mary L. Fonseca, State Senator, Fall River
Ann C. Gannett, State Representative, Wayland
Frank J. Matrango, State Representative, North Adams
William L. Saltonstall, State Senator, Manchester

Staff
Ronald J. Fitzgerald, Director of Research
Ronald B. Jackson, Associate Director of Research
Allan S. Hartman, Associate Director of Research
Joan Fitzgerald, Administrative Assistant
Mary S. Gammons, Head Clerk
Joy A. Fitzgerald, Secretary
Janet F. Flemming, Secretary
Catherine O'Neill, Secretary

Copyright © 1975 by the Massachusetts Advisory Council on Education
Commonwealth of Massachusetts, 182 Tremont Street, Boston, MA 02111

Reproduction of the whole or any substantial part of the contents of this report requires prior written permission of the Advisory Council on Education. It is the current policy of the Council to give such permission for non-commercial reproduction.

MANAGEMENT STANDARDS STUDY ADVISORY COMMITTEE

Jack Childress, (Chairman), Vice President, Educational Testing Service

Thomas Anderson, Remedial Education Instructor, South Middlesex Opportunity Council

Charles Baker, Student, Auburn High School

Michael J. Daly, Massachusetts State Representative, House Chairman, Committee on Education

Harry Daniels, Mathematics Coordinator, Newburyport High School

Joseph Fellows, Counselor, Holliston High School

Henry Fitzgerald, Student, Waltham High School

Ronald J. Fitzgerald, Director of Research, Massachusetts Advisory Council on Education

Leon M. Goodman, Training Supervisor, New England Telephone Learning Center

Richard Goodman, Superintendent of Schools, Wellesley Public Schools

Bertram H. Holland, Executive Secretary-Treasurer, Massachusetts Secondary School Principals' Association

Ronald B. Jackson, Associate Director of Research, Massachusetts Advisory Council on Education

Bedros Kamitian, Principal, Holliston High School

Ara Karakashian, Principal, Reading Memorial High School

John E. Kearney, Director, Regional Centers, Massachusetts Department of Education

Arthur J. Kenney, Principal, North Reading High School

Elaine Kistiakowsky, League of Women Voters of Massachusetts, Cambridge, Member of Massachusetts Advisory Council on Education

Richard J. Lavin, Executive Director, Merrimack Education Center

John Losco, Principal, Revere High School

Alfred A. Maffeo, Associate Executive Director, Massachusetts Association of School Committees

John F. Monbouquette, Principal, Norwood High School

Thelma Moss, Executive Director, Massachusetts Experimental School System

John O'Bryant, Director, Dimock Community Health Center

Robert O'Donnell, Assistant Director of Evaluation, New England Association of Schools and Colleges

Mary Ellen Preusser, Parent, Cambridge, Massachusetts

Edward Sanders-bey, Principal Planner, Massachusetts Department of Education

Harold Shively, President, Bunker Hill Community College

Richard J. Wakeley, Headmaster, Dover-Sherborn High School

Harry Walen, Guidance Counselor, Needham High School

Robert Watson, Director, Curriculum Services, Department of Education

Neal Yanofsky, Student, Newton South High School

CONTENTS

Massachusetts Advisory Council on Education	ii
Management Standards Study Advisory Committee	iii
Foreword by Ronald J. Fitzgerald	vii
Preface by Robert Kates and John Fremer	ix
PART I: APPROACHES AND STANDARDS	1
1. Introduction	3
2. Graduation Requirements in Massachusetts	7
3. Experiential Learning: Current Practices and Needs	11
4. Establishing Management Standards	19
5. Setting Performance Standards	29
PART II: ASSESSING STUDENT COMPETENCE	37
6. Overview of Assessment Methods	39
7. Testimony	45
8. Committee Review	47
9. Performance Assessment	53
10. Product Assessment	61
11. Objective and Essay Examinations	67
12. Looking Back and Looking Around	85
Appendix A: Sample Interview Schedule Initial Needs Assessment Survey	87
Appendix B: New York External High School Diploma Program	91
Appendix C: Dynamy High School Internship Program (Worcester Public High Schools)	93
Appendix D: Executive High School Internship Program	94
Appendix E: The Junior Achievement Program	95
Appendix F: The Home Base School, Watertown, Mass.	97
Appendix G: Metropolitan Cultural Alliance of Greater Boston	99
Appendix H: Sample Documents for a Performance Test	101
Selected MACE Studies	108

FOREWORD

We exist in an era when many different learning options are being developed for students. More and more recognition is being given to the fact that different students sometimes learn best in different environments. Therefore, many high schools are offering more nontraditional and out-of-school experiences to students. Within this context, there are many ways other than written examinations through which students can demonstrate their learning and their right to a diploma. As a high school faces these realities, one major need becomes paramount--THE NEED TO MAINTAIN QUALITY CONTROL BY ENSURING THAT APPROPRIATE PROCESSES ARE USED TO DETERMINE WHAT LEARNING HAS OCCURRED IN RELATION TO WHATEVER LEARNING GOALS WERE BEING PURSUED. Without such processes we cannot be sure that a particular learning option is either effective or ineffective. We can only be sure that it exists. That is not enough.

This Handbook provides educational managers and citizens with a resource guide to important procedures which can measure effectiveness of existing programs for learning. All of the actions suggested herein might not be appropriate for any one school. However, the general management principles promoted in the Handbook are important to every community:

1. A high school or any school community should define clearly the learning goals which it will require of and/or offer to individual students. State goals should serve as a general starting point for local efforts in this regard.
2. To the degree possible, the school should establish clear and realistic standards for individual performance in relation to both common and individualized learning goals.
3. The school community should match methods of assessing individual performance to the learning environment and to the type of learning under consideration.
4. The school should use assessment information both to guide students toward higher levels of achievement and to guide changes in program or learning options. Again, information from statewide assessment can be useful for local comparison and planning.

By observing these principles, educational managers can make a major contribution to the quality of student learning.

The Advisory Council on Education expresses its deep appreciation to the study team, advisory committee members, and the other citizens whose efforts made this Handbook possible. We also thank the Massachusetts Association of Secondary School Principals for the leadership which stimulated this project and which can now become the vehicle for implementing its recommendations.

Ronald J. Fitzgerald
Director of Research
for the
Advisory Council on Education

PREFACE

Evaluation is not a new idea to students and teachers: The former are accustomed to taking tests, receiving grades, seeing averages appear on reports to their parents; the latter to preparing and grading examinations, converting the combined results over a semester into a single grade, recording and reporting the information through a formal administrative procedure.

In a typical classroom course, many evaluation techniques are used ranging from objective and essay examinations to projects, term papers, or models. The final grade in a course reflects student performance and teacher judgment over the semester or quarter. A teacher is aided in this evaluation by the ability to compare the performance of one student with others in the class who have shared the common learning and evaluation experience.

When students move into a learning situation outside the classroom, many of the same evaluation techniques may be used. However, since the experience may be unique for the student and the evaluation may be performed by a number of supervisors and an equal number of agencies or organizations, understanding and agreement on the approaches to be used are particularly important. A rational system of evaluation understood by all parties must be used to monitor student progress and measure program effectiveness. The members of the MACE study staff have designed this Handbook to assist a school in developing such a system, drawing on the talents of teachers, administrators, and members of the community.

The Handbook presents for consideration by the Board of Education and school boards in local communities recommendations for dealing with the need to provide better information on graduation requirements at the state level and more specific statements of goals at the community level. The needs which these recommendations address were made evident to the study staff in a series of town meetings attended by more than 250 parents, students, teachers, and administrators across the state. While the initial presentations to the town meetings dealt with the process and techniques of evaluation, the participants voiced their concern that unless the community and the state had a clear understanding of what students should know when they complete high school, even the best evaluation methods may be wasted when applied to inadequate or inappropriate goals and objectives.

In the Handbook, we have reviewed the context within which nontraditional learning programs are presently being carried out and we have described strategies for combining the professional expertise of school staff with the contributions of parents and other community members. High schools in the state have begun a variety of nontraditional approaches and are improving the quality of these programs. We did not attempt to compile a listing of the various programs. However, in the appendix there are descriptions of a few programs which provide a variety of experiences or a particular credentialing approach that could be used by a high school. Administrators of each program described can provide additional information to interested individuals.

Time and resources did not permit the study staff to develop suggested courses of study or graduation requirements. The intensity of discussion in town meetings and other gatherings indicated a deep concern by citizens and educators and should alert responsible officials in the state to the need for greater leadership in this area.

We want to acknowledge the contributions to this project made by our colleagues at the College Entrance Examination Board and Educational Testing Service, particularly the following:

- A. Graham Down, currently Executive Director of the Council for Basic Education and formerly on the staff of the College Board, director of the MMSPA pilot project

- George Elford and Kathleen Rice, our coauthors and collaborators on the project. Mr. Elford is the author of Chapter 4 and a major contributor to Chapter 3. Ms. Rice is the author of Chapters 1, 2, and 3.

- Matt Sherman, author of Chapter 8 and a contributor to the discussion of goals and objectives in Chapter 4

- Michael Zieky, provider of the material in Chapter 5 on methods of setting standards

- Eleanor Horne, provider of the material in Chapter 11 on methods of finding information about tests

- Paul Campbell, Ernie Kimmel, Jack Osander, Amiel Sharon, and John Valley, reviewers and critics of our plans, methods, and language

- Debra Kaufman, administrative assistant for the project

- Shirley Duby, Peggy Parell, Donna Fleig, Vi Winslow, and Judy Kelly secretaries to the project staff

- Nat Hartshorne, editor of the final version of the Handbook, and Estelle Bartels, Editorial Assistant

Robert Kates
John Fremer
Project Directors

PART I: APPROACHES AND STANDARDS

CHAPTER 1

INTRODUCTION

Today, high school students throughout Massachusetts are participating in many alternative modes of education and evaluation that challenge the administrative procedures and credit requirements of the traditional American high school. Long-established programs continue to provide the basic educational experience for most Massachusetts secondary school students, but there is an increasing tendency to integrate nontraditional learning activities into students' programs. The traditional resources of the secondary school--teachers, classrooms, laboratories, and so on--are not being overlooked; indeed they are being heavily used, often in new and different ways. Considerable use is being made, though, of such community resources as local industries, governmental agencies, and cultural institutions. School and nonschool resources are often combined in alternative or nontraditional modes of education, including such activities as independent study and community-based education, which are sometimes referred to as "experiential learning," "action-learning," or "MAX-ED." Nontraditional methods of evaluation include such techniques as credit-by-examination, committee review, and product assessment.

Nontraditional modes of education such as community-based programs have great potential as a way of both meeting traditional high school goals and reducing the barriers that exist between the adolescent and the broader community. For example, students who have participated in executive internship programs in the business and service industries have had the opportunity to try on occupational roles, to practice independent and responsible action as well as cooperation, and to benefit in general from realistic contact with the adult world. At the same time, however, educational activities that take place outside the high school setting pose special problems: How are such experiences to be described, reported, and evaluated? Should credit be awarded for such experience? If so, how much and how should it be granted? Likewise, nontraditional assessment techniques pose special problems: What techniques should be used to measure what experiences? Which student populations can benefit from which techniques?

Today, as nontraditional learning programs continue to proliferate, it is more crucial than ever before to find ways of dealing with these problems. This Handbook of Management Standards for Awarding High School Diplomas represents one step in that direction.

The Origins of the Handbook

As early as 1970, the Massachusetts Secondary School Principals Association (MSSPA) recognized the promises as well as the problems involved in nontraditional modes of education and evaluation. This group identified and articulated both the need for quality control in alternative learning programs as well as the need for more and varied ways to certify learning accomplishments.

The early efforts of the MSSPA Executive Board resulted in the establishment of a Committee on Diploma Requirements. During the spring of 1974, upon the recommendation of this Committee, MSSPA conducted a project with the funding and staff support of the College Entrance Examination Board. The project was designed to explore ways of establishing a secondary school-level credit-by-examination program, which was one of the recommendations of the National Committee on the Reform of Secondary Education.

The Massachusetts Advisory Council on Education (MACE) shared the principals' interest in nontraditional assessment techniques and their concern for quality high school education in both its traditional and nontraditional forms. During the summer of 1974, the members of MACE contracted with the College Entrance Examination Board and Educational Testing Service to undertake a project they called Management Standards for Awarding High School Diplomas.

A Dual-Purpose Project

The two-fold purpose of the project was (a) to produce a handbook to assist secondary school administrators in the quality control of alternative learning experiences and in the development and application of alternative methods of evaluating such experiences; and (b) to generate statewide discussion and exploration of the various issues related to nontraditional high school education.

Drawn, in part, from a study of innovative programs--particularly those in Massachusetts, New York, and Oregon--this Handbook provides specific recommendations on management, evaluation, and resources for certifying learning accomplishments that may qualify a student to receive a high school diploma. Since each school district must adopt and/or devise programs that best suit the needs and preferences of its own constituency, the Handbook has also been designed to provide the background information on nontraditional education and evaluation needed to make decisions about individual programs.

In order to generate a statewide discussion of the various issues involved in nontraditional education and solicit community support and direction, an Advisory Committee was formed composed of members of the business and educational communities, parents, students, and legislators. The function of the Committee was to react to the plans for the study, advise on the development of goals, suggest priorities for the development of evaluation measures, and comment on the Handbook prior to its publication.

To solicit opinion and reactions, presentations of the nature and purpose of the project were made to major professional organizations. To focus on student involvement, a Student Advisory Council was formed by the Department of Education.

To generate discussion at the local level, town meetings were held throughout Massachusetts in cooperation with the Regional Educational Centers of the Department of Education. Designed for parents, students, faculty, administrators, and other interested members of the community at large, the meetings provided an opportunity to disseminate information on the nature of the project and its possible outcomes and record the reactions.

As the study progressed and as more and more people voiced their opinions, two issues of central concern emerged:

- There seemed to be widespread support for the idea that some broad goals of education could be met through experiential or community-based learning provided effective measures of evaluation were used to assess the actual learning outcomes.
- Discussions about graduation requirements, standards of performance, community expectations, and the meaning of the high school diploma revealed limited knowledge or understanding of what current graduation requirements are and strong support for the idea that, as a result of the high school experience, students should be able to demonstrate mastery of basic competencies. This latter idea is also reflected in the educational goals published by the Board of Education in 1971, which could provide the focus for some degree of competency-based education in the Commonwealth.

The abundant interest in and concern with current graduation requirements and competency-based education justified further exploration and discussion of these two issues. In time, it became apparent that the best response to these concerns should be study recommendations. Thus, the project staff formulated three recommendations and presented them for review by the Advisory Committee. The recommendations, outlined on this page, are reflective of the central concerns raised in discussion across the state.

1. THE BOARD OF EDUCATION SHOULD PUBLISH A RESOURCE DOCUMENT OUTLINING THE LAWS AND REGULATIONS AFFECTING GRADUATION REQUIREMENTS AND PROVIDING INFORMATION ON THE OPTIONS THAT CAN AND/OR SHOULD BE EXERCISED BY COMMUNITIES AT THE PRESENT TIME.

2. COMMUNITIES, THROUGH THEIR SCHOOL COMMITTEES, SHOULD ADOPT MINIMUM STANDARDS, BASED ON COMMUNITY GOALS, FOR THE DEVELOPMENT OF STUDENT COMPETENCIES IN THE BASIC AREAS OF READING, WRITTEN AND ORAL COMMUNICATION, COMPUTATIONAL SKILLS, SOCIAL SCIENCE-HISTORY, PHYSICAL EDUCATION, AND SUCH OTHER AREAS AS THE COMMUNITY DETERMINES TO BE ESSENTIAL.

3. FOR THE GUIDANCE OF LOCAL SCHOOL DISTRICTS, THE BOARD OF EDUCATION SHOULD PROVIDE SAMPLE INDICATORS OF PERFORMANCE RELATED TO THE EDUCATIONAL GOALS ADOPTED BY THE BOARD IN 1971.

To provide background information for the first recommendation, the project staff undertook a preliminary investigation to determine current diploma requirements, current practices, and related legislation in Massachusetts. A summary of the results of this investigation is included in Chapter 2.

CHAPTER 2

GRADUATION REQUIREMENTS IN MASSACHUSETTS

As a result of the statewide discussion of nontraditional education described in Chapter 1, attention was drawn to the following questions: What should a student be able to know and do as a result of the high school experience? What should the diploma mean? What are current graduation requirements and what should they be?

Discussion revealed that many people are not aware of or misunderstand current graduation requirements. Since an understanding of current requirements is a prerequisite to any discussion of what might or should be done to further the mission of the high school, a preliminary investigation, reported below, was undertaken. The report of this investigation, called for in Recommendation 1, appears on the following pages. The concerns about what students should know when they leave school prompted an investigation of competency-based education. This led to the formulation of Recommendations 2 and 3 and the general discussion that concludes this chapter.

What the Law Mandates

In order to determine the existing pattern of high school graduation requirements in Massachusetts, the project staff reviewed the General Laws relating to education and examined courses of study in 65 high schools.

High school graduation requirements are determined by local school districts under the provisions of Chapter 71, Sections 1, 2, and 3 of the General Laws relating to education. According to Massachusetts law, school districts are required to maintain a high school that offers one or more courses of study at least four years in length and that must offer a variety of subjects, including English, mathematics, science, and personal and skill development. However, the law mandates ONLY that a student take two subjects: U.S. history and civics and physical education.

The law provides local school districts with great flexibility in the determination of subject areas and the minimum number of credits required for graduation; it does not in any way specify the extent to which alternative methods of education, evaluation, and certification may or may not be implemented. The flexibility inherent in the current law is viewed as highly desirable, and no recommendations for change in the law are suggested.

Current Practices

A marked diversity exists among Massachusetts high schools both in terms of subjects to be taken and the minimum number of credits required for graduation. Inspection of the various systems, interpreted in terms of the

Carnegie Unit*, revealed that the minimum credit requirements for the Massachusetts high school diploma range from 72 to 100 credits and from 16 to 20 units.

In addition to U.S. history and civics and physical education, which are required by law, all 65 schools surveyed require four years of English. Mathematics and science are required by approximately 76 percent of the sample; social studies (in addition to U.S. history) is required by 54 percent, and health, in addition to physical education, is required by 20 percent of the sample. Only 25 percent of the schools explicitly state that required courses, including physical education must be passed, while four require that certain courses must be taken but not necessarily passed. At a number of schools, a student's attendance record counts toward graduation, and most schools indicate a minimum residency of at least one semester in order to qualify for the diploma. Approximately one-third of the sample indicated that activities such as work-study, independent study, and make-up work must have prior approval so that credit accumulated by such experiences may be applicable to graduation requirements.

Early graduation requirements, specified by more than half the schools, include completion of regular course and credit requirements (except physical education, which may be waived), demonstration of social maturity, above-average school work, and a conference between school and family to ascertain whether the student has an educationally appropriate substitute for the senior year. Educational Research Service has recently conducted a survey of regulations (1) pertaining to early graduation from high schools in selected states. The report examines the current status of early high school graduation both in terms of the school policies and the students who elect this alternative.

Competency as a Diploma Requirement

Under Public Law 572 of the laws of 1975, the General Laws relating to education were modified. Section 1G of Chapter 15 states that the Board of Education shall establish "minimum educational standards for all courses public schools require students to take" (2). Despite the interest in competency-based education, initital efforts to determine "minimum educational standards" appear to have dealt more with reinterpretations of the existing curriculum, the physical plant, and personnel and organizational matters than with minimum competencies expected from recipients of the high school diploma. However, in a few instances, school districts have introduced requirements for students to demonstrate competency in writing and computational skills, and in at least one district, competency in biology.

In competency-based education, it is the specific things students can do rather than time spent in courses that is recognized. The programs tend to focus on the outcome of learning with more emphasis on performance than on conceptualization--more doing than knowing--with standards based on performance criteria instead of norms. Traditional educational quality controls

---

* For purposes of this study, the credit and point systems of the 65 sample high schools were translated into a unit system in which five credits or points are equal to one unit or one year of full-time study (Carnegie Unit).

have focused on such inputs as a fully equipped school, libraries, and a qualified faculty, the assumption being that the quality of learning is directly related to teaching conditions. On the other hand, the assumption in competency-based education is that quality control is better achieved by focusing on learning rather than teaching. The requirements involve the demonstration of skills and knowledge called "competencies." Of each learning experience, the following questions are asked:

1. What skills or knowledge are to be learned and how do they relate to program and school goals?

2. How well are they learned?

3. Under what circumstances are they to be learned?

4. What is the evidence that will be presented to indicate that they were learned?

Competency-based education is operative today throughout Oregon where the Board of Education in 1972 adopted graduation requirements that spell out three general areas in which students must demonstrate minimum competency (3). The three areas and their meanings are:

1. Personal Development: Each student shall demonstrate competencies necessary to:

   a) Read, listen, analyze, speak, and write.
   b) Compute, using the basic processes.
   c) Understand basic scientific and technological processes.
   d) Develop and maintain a healthy mind and body.
   e) Develop and maintain the role of a lifelong learner.

2. Social Responsibility: Each student shall demonstrate the competencies required to function effectively and responsibly:

   a) As a citizen in the community, state, and nation.
   b) As a citizen in interaction with his or her environment.
   c) As a citizen on the streets and highways.
   d) As a consumer of goods and services.

3. Career Development: Each student shall demonstrate competencies required to function effectively within a career cluster or broad range of occupations.

In order to help schools implement competency programs and evaluate the extent to which the broad goals are met, the Oregon Department of Education has provided models and guidelines which translate the general goals into intermediate goals (program goals) and then to more specific goals (course goals). One school district in Oregon lists four program goals and 20 course goals (competencies) that constitute the social responsibility goal that "each

student shall demonstrate the competencies required to function effectively and responsibly as a citizen on the streets and highways." An example of a program goal is: "...students will be able to function as responsible individuals when operating automobiles." One of the competencies (course goals) is: "...the student will be able to demonstrate minimum performance capabilities required for entry into the highway traffic system as a motor vehicle operator." For each course goal the state provides sample indicators of the desired performance. In this case, three sample performance indicators are given. One is: "...given an opportunity, the student will demonstrate his ability to change a tire."

Emphasis is placed upon the provision of as many performance indicators as possible since, due to differences in individuals, competence is demonstrated in many ways. For example, a further indicator developed by one of Oregon's districts is that the student would pass a written driver's examination or cite the proper highway conduct for pedestrians, bicyclists, and motorists.

Competency-based education used in conjunction with the traditional course/credit system is a new and essentially untested approach to high school education. Thus, there are some important problems yet to be solved. One of these is that performance is often complex and difficult to define. Another is that standards are elusive. If competence is defined in vague terms without specific assessment guidelines, the highly verbal student who has the quick and glib response to a teacher's questions may be at a special advantage; if defined in terms that are too specific, the long-valued objective of schools of providing a broad educational experience for students may be subverted.

These are the kinds of problems that are addressed in the remaining chapters of this Handbook, which provide models and resources for planning, standard setting, and assessment. In this sense, the various chapters represent steps toward solving some problems inherent in nontraditional education in general and in competency-based education in particular.

REFERENCES

1. Stemnock, S.K. Early graduation from high school: Policy and practice. Arlington, Virginia: Educational Research Service, Inc. 1975.

2. The Commonwealth of Massachusetts, Department of Education. General Laws Relating to Education. Boston, Massachusetts: The Commonwealth of Massachusetts, Department of Education, 1970.

3. Oregon Board of Education. Oregon graduation requirements: guidelines for record-keeping procedures and sample forms. Salem, Oregon: Oregon Board of Education.

CHAPTER 3

EXPERIENTIAL LEARNING: CURRENT PRACTICES AND NEEDS

To describe what is meant by nontraditional or experiential approaches to learning, it is useful to recall the distinction that people commonly make between schooling and education. People have long recognized that a person can acquire an education or learning without going to school. In fact, many would suggest that there are some things that can best be learned outside of schools. Traditional learning refers here to what is usually thought of as "school learning." Nontraditional learning, even when it is arranged for by the school, recognizes that education or learning goes on continually in many settings. As the report of the 1974 Panel on Youth (1), chaired by James S. Coleman, has pointed out:

> We instruct one another informally and learn something new as we raise a family, earn a living, minister a church, nurse the ill, encounter friends, govern a group or communicate through a mass medium.

The terms "experiential learning" and "action learning" are also used along with "nontraditional" and "nonformal" learning. The use of terms such as "experiential" and "action" to describe the learning which takes place outside of the regular school program is based on the emphasis in these programs on experience and action. This is not to suggest that all classroom learning is devoid of any experience or action.

Definition of Traditional and Nontraditional Approaches

According to the Coleman panel (1), in the school "teaching becomes specialized work and learning takes the form of systematic participation that is socially separated in a student role and assembled in a particular organizational place."

According to William G. Spady (2), the school's role as an agent of the larger society entails five distinct but interdependent functions: instruction, socialization, custody-control, evaluation-certification, and selection. These roles imposed by the larger society on the school produce certain constraints that nontraditional approaches must take into account.

As the term implies, "nontraditional" is defined in reference to the usual approaches to learning. Along with nontraditional, the term "nonformal" learning is sometimes used to distinguish this learning from formal or traditional learning. In secondary schools, traditional learning takes place in or in connection with scheduled classes. The classes meet with a teacher in the school building as a part of a course of study leading to a diploma. Textbooks, teacher-made tests, written homework, and assigned reading are among the usual

fixtures of traditional learning. This learning is directed chiefly toward cognitive and occasionally toward a narrow range of psycho-motor objectives. The affective objectives usually turn out, in fact, to be related to the custody-control function of the school, as defined by Spady. According to Spady (2), this custody-control function plays an important role in traditional high school learning:

> Concern for custody-control seems to have pervaded the school to such an extent that it also impinges on the school's evaluation-certification function. This "contamination" is apparent at both the classroom and school system levels. In the classroom, for example, teachers control the evaluation process and commonly dispense grades not only on the basis of student performance and effort but on punctuality, attendance, "attitude," and deportment as well. At the school system level formal certification and the awarding of diplomas depend as much on students' exposure to twelve years of schooling as on the skills they acquire.

The instructional function, as undertaken in the traditional approach to learning, might best be described by citing current practices as revealed in a 1972 national sample study of 18,000 high school seniors. According to this study (3):

> In an eight-item list, "listening to the teacher's lecture" was marked most often (by 83 percent) as having been used fairly often or frequently in their senior-year courses. The following three methods were also employed often with a large proportion of the students: "Writing essays, themes, poetry, or stories" (64 percent), "participating in student-centered discussions" (60 percent), and "working on a project or in a laboratory (48 percent). On the other hand, only small proportions of seniors reported that their courses employed the following four methods fairly often or frequently: "Having individualized instruction (small groups or one-to-one session with a teacher)" (24 percent), "using teaching machines or computer-assisted instruction" (12 percent), "going on field trips" (11 percent), and "watching television lectures" (6 percent).

While the majority of the seniors surveyed admitted that their own poor study habits had interfered with their learning, they also complained that the school didn't offer the courses they wanted to take. Two-thirds of the seniors agreed that the "school does not offer practical work experience." The survey, however, also showed that over three-fourths of the 1972 class worked during their senior year. Some 44 percent of the employed boys and 29 percent of the employed girls worked at least

20 hours per week. Just over one-third of those employed felt that their jobs would lead to work they would like to do in the future.

Nontraditional learning, for present purposes, describes those learning experiences that take place outside a regularly scheduled high school class or course. In this usage, nontraditional learning would range from independent study or special project work within the school to work-study or action-learning programs that are carried on principally away from the school.

The Need for Nontraditional Approaches

Concern for nontraditional approaches to learning is ultimately one for students. The following list attempts to describe the kinds of students for whom, at some point in their education, the traditional approaches may fall short of meeting their needs; these are the students for whose benefit nontraditional approaches should be devised. These types of students are not listed in any particular order.

- The artistically oriented: With the advent of the comprehensive high school and the comparative decline of specialized schools, artistically oriented students often require learning opportunities that go beyond the range of offerings in their school's art, drama, or music departments. Art museums, conservatories, studios, or community theaters might better serve their needs.
- The skill-oriented: The school's shops and vocational educational programs might fall short of the needs of many skill-oriented students. Auto or TV repair shops, industrial training programs, larger vocational education centers, or expanded apprenticeship programs might better meet their needs.
- The scientific-minded: Students who have an unusually keen interest in science might need to pursue projects and studies in libraries, museums, zoos, and laboratories within the community.
- Potential dropouts (or "top-outs"): Academically poor and restless students might profit from some change from a curriculum which demands too much of them and in which they have no interest. Bright and restless students for whom the curriculum is an undemanding bore might also profit from some alternative possibilities not possible within the school's course offerings.
- The special student: Students with physical handicaps or learning disabilities would benefit from out-of-school opportunities for learning in areas which are compatible with their special needs and interests.
- Youth in need of acceptable work habits and a sense of responsibility: Some programs are now offered to 14- or 15-year-old students who are unlikely to complete the regular high school program and need special help and counseling in developing acceptable work attitudes and habits. For these students, a work-study program is not for teaching specific skills; it is designed to foster habits and attitudes, such as dependability, punctuality, acceptance of construc-

tive criticism, and respect for authority.

- Career explorers: For some students, career exploration might provide the basis for a work-study program that supplements the career-awareness programs offered within the school.
- Career beginners: Students who have the opportunity while in school to begin a career path should be supported by special work-study and counseling programs.
- Independent students: Some students might work better on their own in some areas in which they are now held back by the class requirements. Independent study and the means for fairly crediting such study should be made available.
- Students with special learning opportunities: Students who learn language or cultural history as a result of family or group travel and students with other special opportunities should have the means to pursue these opportunities and to receive credit for the consequent learnings.
- The service-oriented: A number of "action-learning" programs already tap the interest and abilities of high school students for community service in hospitals, government agencies, day-care centers, and so on. Schools also use students' interests and skills by having them participate in tutorial programs. The incidental learning from these experiences along with related study by the student are deemed worthy of credit.

The Coleman panel identified two needs not associated with any particular subgroup but attributed to all high school students. The panelists saw a need for most students to overcome the school-imposed segregation of the young from society and the world of work and the segregation of any one group of students from other age groups. Clearly, one characteristic of traditional learning is that in school students spend most of their time with other students of their own age. The objection here to the traditional approach to high school learning is related, in Spady's terms, more to socialization -- the transition from youth to adulthood -- than to instruction. The Coleman panel members allowed that "at their best, schools equip students with cognitive and non-cognitive skills relevant to their occupational future" and "with knowledge of some portion of civilization's cultural heritage." According to the Coleman panel (1), however,

> they do not provide extensive opportunity for managing one's affairs, they seldom encourage intense concentration on a single activity, and they are inappropriate settings for nearly all objectives involving responsibilities that affect others. Insofar as these other objectives are important for the transition to adulthood, and we believe they are, schools act to retard youth in this transition, by monopolizing their time for the narrow objectives that schools have.

The panel elsewhere explained (1):

> the essential difficulty of schools in handling activities other than academic learning is the position of the youth within the school. He is a dependent, and the school is

> responsible for shepherding his development. Yet if youth is to develop in certain ways involving responsibility and decision-making, then the responsibility and dependency are in the wrong place. To reorganize a school in such a way that young persons have responsibility and authority appears extremely difficult, because such reorganization is incompatible with the basic custodial function of the school.

Three other reports on secondary education that have currently been either published or publicly reviewed have called for new approaches to high school education. An unpublished study by the U.S. Department of Health, Education, and Welfare's National Panel on High Schools and Adolescent Education chaired by John Henry Martin underscored the lack of contact between youth with adults other than teachers or parents. According to the Martin panel (4), two of the major objectives of education have failed at the secondary level:

> Education for citizenship has not practiced what it preached since "student participation in school management has been kept at a safe distance from the real discharges of responsibility." And the favorable position of vocational education in schools "is enjoyed more out of expectation than from performance" with youth employment due more to opportunity than to preparation.

The Martin panel recommended that the comprehensive high school be replaced with comprehensive education, with youth and adults working together in community art centers and community career education centers. According to this panel, required attendance at high schools should be limited to a maximum of four hours per day, with the basic role of the high school limited to essential skills. Earlier, the Kettering Foundation's panel report, entitled The Reform of Secondary Education (5), called for the compulsory attendance age to be lowered to 14. In this panel's words, "If the high school is not to be a custodial institution, the state must not force adolescents to attend. Earlier maturity requires an option of earlier departure from the restraints of formal schooling." (These reports agree that adolescents are maturing much more rapidly today, with puberty occurring two years earlier than it did two generations ago.) The Kettering report devoted a full chapter and four recommendations to the need for "nonformal sources of secondary education." More recently, a task force on secondary education of the National Association of Secondary School Principals (NASSP) disagreed with the above recommendation on ending compulsory schooling at age 14. This NASSP panel (6) instead called for a new role for the high school in which the school would serve as a "planner and broker" for learning but not necessarily as the dispenser of education for students after age 14. In this new arrangement, students would be free to come and go within the school as they please through age 17. They could take time off, learn outside of school, and come back if they wish. The term "dropout" would be applied only to one who moved in no constructive direction, despite the school's new scope. The dropout would become the responsibility of other youth-serving agencies.

In general, these recent reports all are in agreement that there must be some relief from the custody-control function that burdens the high school in such dramatic ways today. Nontraditional learning does not necessarily imply abandonment of the custody-control function of the school. At present, for example, action-learning programs generally make special provisions for this aspect of the program, operating on a "behave or get out" basis.

Current Practices in Nontraditional Programs

To identify current practices in nontraditional approaches to learning, the project staff in November 1974 surveyed the twenty-plus action-learning programs identified in the NASSP's brief report entitled Twenty-five Action-Learning Schools (7). According to this survey, students in these programs were receiving credit for all forms of paid work and/or volunteer service. While the formal mechanisms through which a student may work or volunteer his time varied widely, most experience was tied to classroom work, connected to an academic discipline, or linked to post-high school career exploration. While all schools utilized some control device to assure that the student's experience had educational merit, the prevailing viewpoint held that all experience has educational merit because work or service provides the youngster with the opportunity to be responsible, to gain competence in a working situation, to give service to others, to interact across age groups, to learn about group action and the adult work situation. Thus, students were assisting in medical research in California, organizing search and rescue teams in Oregon, working with the aged or the retarded in Vermont, and teaching language or math skills to their colleagues in Wisconsin.

All schools surveyed indicated that objectives, the student's role, and performance criteria were set before the student reported to his or her assignment.

While some schools appeared to have systematic and comprehensive evaluation procedures, all discussed various problems inherent in any kind of evaluation. The majority of the schools suggested that on-site visitation by a school coordinator and periodic evaluation were necessary for quality control and fair evaluation.

All schools translated work or service experience into credit. The amount of credit the student received was determined by the hours he or she had put in. Half of the schools surveyed indicated that the student's experience was translated into an academic discipline and then recorded on the transcript. The other half simply recorded the name of the program in which the student participated. This latter group strongly espoused the idea that credit for all learning experience in the classroom or off campus be determined by the same criteria.

The school representatives generally agreed that fair and consistent evaluation was very difficult to achieve; however, they added that fair and consistent evaluation was also very difficult to achieve in traditional programs. Quality-control problems generally arose from 1) student's motivation -- to get out of class, to make a buck; 2) agency motivation -- to use the student in menial tasks exclusively as cheap or free labor; 3) the lack of adequate criteria in certain jobs and service areas. The consensus of opinion among those surveyed was that credits for action-learning programs were no more deceptive than credits in traditional courses.

Figure 1

Assessment Models for Nontraditional Learning Activities

Learning Activity	Description	Suggested Assessment Models				
		Testimony	Committee Review	Performance Assess.	Product Assess.	Objective and Essay Exam
Work Study or Cooperative Education	Work related to learning skills or attitudes and habits or career exploration, accompanied by study	X	X			
Action Learning	Community service in hospitals, day care centers, agencies, or tutoring	X	X			
Independent Learning (course-related)	Study or research projects in or away from school		X			X
Independent Learning (not course-related)	Same as above, also travel and special experience learning	X	X			
Artistic Performance	Performance in dance, drama, or music			X		
Art or Craft Product	Products or works in art, ceramics, sculpture, etc.				X	
Vocational Arts-Performance	Work in auto repair, TV repair, etc.			X		
Vocational Arts-Product	Products in carpentry, masonry, electrical wiring, cooking, serving		X		X	
Literary Works	Essays, poetry, short stories, etc.		X		X	

For example, in an algebra course, a student receives a D, obtains credit, and is certified to have "learned" algebra; and yet the student himself will admit that he has not, in fact, learned algebra.

In addition to surveying the programs in the NASSP's report, the project staff identified a number of unique and promising nontraditional programs in high school education. Descriptions of these programs and the methods used to evaluate them can be found in Appendixes B-G.

Assessment in Nontraditional Approaches

Alternative forms of learning require, to some degree, alternative approaches to the assessment and certification of learning. Teacher-made tests, term papers and other written works, classroom recitation, lab projects, products produced in home economics, art, and shop are among the evaluation-certification tools in the traditional program. While a number of these approaches can be translated for use in nontraditional programs, as the project survey made clear, the problems of evaluation in nontraditional programs merit special attention. In Figure 1, the various categories of learning experiences associated with nontraditional approaches are related to possible models of assessment that could readily be used in evaluating these learning experiences. Admittedly, in several instances, both the learning experiences and the assessment models are found in traditional programs as well. Subsequent chapters of this Handbook will examine these assessment models in detail.

REFERENCES

1. Youth, transition to adulthood. Report of the Panel on Youth of the President's Science Advisory Committee, James S. Coleman, Chairman. Chicago: University of Chicago Press, 1974, 76, 142, 146.

2. Spady, W. G. Research on Oregon's competency-based system of high school requirements. Unpublished paper, 1974.

3. Fetters, W. B. A capsule description of high school seniors. Washington, D.C.: U.S. Government Printing Office, 1974, 7.

4. Martin, J. H. Studies agree on changes needed in high schools. Education USA, November 18, 1974, 17, No. 12.

5. National Commission on the Reform of Secondary Education. The reform of secondary education. New York: McGraw-Hill, 1973, 21.

6. NASSP sees new, flexible role for high schools. Education USA, February 17, 1975, 17, No. 24.

7. Twenty-five action-learning schools. Reston, Virginia: National Association of Secondary School Principals, 1974.

CHAPTER 4

ESTABLISHING MANAGEMENT STANDARDS

Innovations such as experiential learning programs will be accepted more readily if standards are agreed upon and established at the outset. Two kinds of standards will be described in this Handbook--performance standards and management standards. A performance standard indicates the level of performance, such as a minimum score on a test, a minimum rating on a rating scale, or a minimum course grade that is required for a certain purpose. For example, schools often require that a student earn a "D" grade before credit is awarded. In most high school courses, the performance standards are determined by the teacher in terms of the quality (or quantity) of work required for an "A", for a "B", and so on. At times, teachers describe these standards in a clear and precise manner; at other times, the standards are barely discernible, even to the teacher.

The setting of performance standards traditionally has been and should remain the responsibility of the local school staff. The faculty can then assume responsibility for the worth of their school's grades and credits. The same responsibility applies to the grades and credits associated with nontraditional as well as traditional learning. Chapter 5 of this Handbook focuses on problems and possible approaches in the setting of performance standards.

Management standards describe what steps are to be taken and what processes are to be followed so that, in the designing and implementing of any programs, necessary decisions are made by qualified people in a manner that provides for consistent quality. For example, suppose a teacher arbitrarily or from force of habit set a passing score of 70 on a 100-item final examination. While this performance standard might prove to be quite appropriate, the process whereby the standard was set, with its arbitrariness or reliance on habit, certainly failed to meet reasonable management standards.

While the term itself might seem new or unfamiliar, management standards have been followed for some time in the evaluation and crediting of student performance in traditional high school courses. Management standards require that student performance standards be determined and that student performance in course work be evaluated by a teacher, who is supervised by a department chairman, and the school principal, all of whom carrry appropriate state certification for their positions. In addition, the state and the regional accrediting agency attest to the adequacy of the staff, programs, and facilities of the school. School accreditation and school personnel certification processes are the chief assurances to the public that reasonable management standards are being maintained in the schools. Because nontraditional programs are new and may reach beyond presently accredited programs and certificated personnel, special attention must be given to establishing new and adequate management standards.

The MACE Handbook and the Change Process

For a number of high schools in Massachusetts as well as in other states, learning outside the classroom and the assessment of this learning would be considered an innovation. For this reason, some consideration of how educational innovation aakes place should prove helpful. Thomas Woods (1) has described the adoption of educational innovations as a five-step process involving: (a) initial awareness of the innovation as a possibility, (b) an interest in this possibility leading to further inquiry, (c) some evaluation of the innovation prior to use (shall we try it?), (d) a trial use in a carefully observed arrangement, and (e) adoption, usually after a successful trial case.

While one would logically assume that educators first identify a need or problem and then look for possible new approaches, Woods noticed that awareness and interest often precede a clear sense of need. In fact, the needs tend to be more clearly identified later in the interest and evaluation stages. Woods also observed that "early users" of an innovation tend to invest more time and care in the trial-use period than do "late adopters " who rely more on the previous experiences of others. In terms of these five stages, this Handbook might well serve for some as a stimulus to an initial awareness and interest in new or nontraditional approaches. It is also meant to assist school staffs in making some evaluation prior to use and in the trial use of these approaches.

Some other information (1, 2) about the change process in education might also have some usefulness here. For example, the success of an innovation depends, in part, on the manner in which it is introduced and, in part, on the characteristics of the innovation itself. Innovations that are seen by the local staff as their own, that enjoy the full support of administrators and policymakers, that are developed with staff involvement both in the diagnosis of the problem and in later decision making, are likely to succeed. The innovation itself would best be reasonable in cost; it should utilize already prepared materials, have readily available supportive services, and prove to be compatible with the existing system. Ventures that clash with vested interests, create job insecurity, or add to staff burdens hold less promise of success.

The particular approach endorsed by this Handbook places a heavy emphasis on the careful identification of the problem or the local needs, with a correspondingly careful statement of the solution requirements. Innovations that are directed toward clearly identified problems or needs tend to be successful. By the same token, many innovations have failed dramatically for want of a clear statement of the need. All too often, innovations are introduced in response to a general set of conditions--a sense of malaise or dissatisfaction with the status quo or a general desire to make improvements or to try something new and different. Sometimes innovations result simply out of a desire to keep up with the Joneses.

The management approach recommended here urges administrators to go beyond general impressions of need and to identify the specific needs or problems. Thus, faced with a promising innovation, administrators should ask specifically: Who needs this? Why? How badly do they need it? Precisely which students have these needs? What would things look like if the needs were met? Will the proposed innovation really meet these needs? If so, in what way and to what extent? Studies of failure in educational innovation have shown these to be the neglected questions.

This Handbook focuses on the high school as the locus for the introduction of new approaches to learning. This is because most innovations in experiential education have taken place in individual schools, which makes the effort less cumbersome and complicated than when a number of schools are involved. However, it is important to note that there are clear advantages in taking a regional, districtwide, or collaborative approach as is being done with the implementation of the recommendations in Chapter 766. Such larger efforts involving a number of schools can draw upon more extensive financial and professional resources. In general, such a large-scale effort would follow the same steps as those outlined here for an individual school. Figure 2 outlines the steps described in this chapter. It is important to note again that these recommended procedures describe only one approach or method for introducing nontraditional learning to a high school.

Suggested Management Standards

The management standards recommended here will take the form of a series of brief prescriptive statements followed by some explanatory comments. These statements provide a checklist to be reviewed. Any set of general management standards will appear somewhat idealized since the concrete realities do not always allow for such a careful step-by-step process. Nevertheless, these standards, which describe one reasonable way of proceeding, should prove useful. The statements assume that an initial awareness of and some interest in nontraditional learning and assessment have already been generated, whether it be from a perusal of this Handbook or from other sources.

The following are recommended management standards:

> The locus of responsibility for implementing any new approaches should be clearly defined.

In most Massachusetts high schools, the key person involved at the school level will be the principal. In a few very large schools, the vice principal in charge of curriculum might be the key administrator. If the experiential program involves only one approach and a limited number of students, the principal might appoint one staff member to direct the program. However, at the outset, the principal plays the key role.

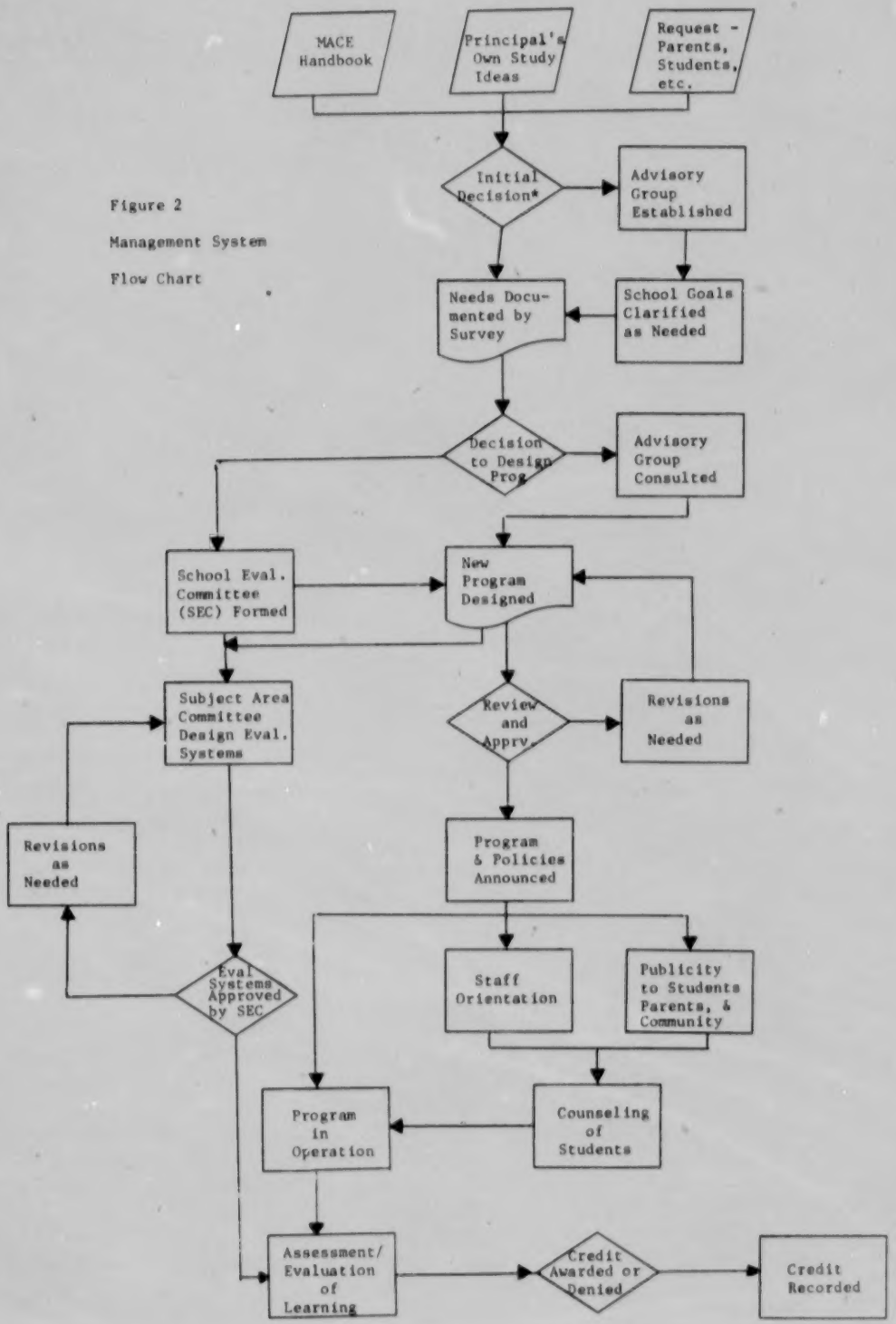
As the educational leader, it is the principal's responsibility to establish certain working definitions. For example, the term "nontraditional learning experience" must be defined. Not all experiences lead to learning or to the kinds of learning to which the school should assign credit. Nontraditional learning, as used here, includes only those experiences that lead to learning outcomes consistent with the goals of the school.

Attention must be given to the school's stated goals, which provide the framework and the justification for any new program. Every high school has its stated goals. These may be perfectly adequate or they may be out of date, too vague, or too narrow. If there is some need to re-examine the stated goals of the school, the principal might organize a community study group to re-formulate these goals. Such a community group would best include parents, teachers, students, business and school committee representatives. This study group should give some consideration to stating the goals of the school insofar as possible in terms of the competencies that the community looks for its graduates to possess. (See the Oregon approach described in Chapter 2.) Statements of school goals, however, should not be very specific, lest they have the effect of trivializing or unduly limiting the mission of the school.

Figure 2

Management System

Flow Chart



* To explore possibilities of nontraditional program

Within the school's curriculum, departmental goals should be more specific, and individual course or program goals should become, in effect, quite specific learning objectives, stated in terms of student outcomes or competencies. The statement of these specific learning objectives in terms of outcomes or competencies is an essential first step in the evaluation or assessment process. Thus, an adequate statement of school goals and later, of course and program goals, is the first step for which the principal is responsible.

Even if the school's goal statement does not need the attention of a study group, the principal should form an advisory group similar in make-up to the group described above to counsel him at every stage in the development and operation of a nontraditional program. If the school already has such an advisory group for all its programs, no new group need be organized for this new program.

A systematic and visible survey of needs should be carried out as the first step.

The very fact that the principal and other staff or students are interested in nontraditional approaches indicates that they sense something of this kind is needed. Interest in new approaches is sometimes produced by a particular crisis or by the discovery of a new opportunity or learning resource that is readily available. The first step is to clarify and make specific the local needs and problems. Informal conversations with staff members, students, and parents should give the principal an initial "feel" for the situation. Following this informal probing, an ad hoc faculty committee including several student representatives would then plan and carry out a needs survey to test several assumptions or impressions formulated by the ad hoc group as to who needs what. This survey might well take place through interviews with a sample of faculty, students, and parents.

A survey here would require contacting not a random sample but one including representatives from deliberately selected categories of faculty, students, and parents. These categories would include those who are presumed to have a particular interest or stake in experiential learning. The list of the types of students in need of these programs (see Chapter 3) provides some direction as to those who are to be included in this purposive sample. For example, at the outset, the committee might assume that a need for new ways of crediting learning might be felt by students learning a foreign language at home, Junior Achievement leaders, students with independent study skills who wish to gain more credits than their present schedule permits, students learning skills at work as auto mechanics, newspaper reporters, electricians, and so on, and students learning from work in the performing arts or fine arts outside the classroom. Once the categories have been identified, a group of students (at least 8 to 10) chosen randomly in each category should be interviewed or questioned about whether they need nontraditional programs and if so, what kinds.

Interviews are to be preferred over questionnaires. If the use of a questionnaire becomes inescapable, the questions themselves should be presented in an open-ended format. Avoid multiple-choice questions here except for background or factual information items. The interview questions should be carefully planned and follow a set but not rigid format proceeding from the general to the specific. (A sample interview schedule is included in Appendix A.)

The outcome of this needs survey should be a report stating the needs related to experiential learning and new modes of assessing learning. In defining a need or problem, it is important not to include in the need or problem statement any mention of a solution. For example, to say that "what this school needs is a credit-by-examination program in auto mechanics" does not accurately describe any need. This statement assumed a need and describes a solution. The important contribution of the needs survey is to identify clearly the needs or problems. Such a clear statement will then lead to a description of the solution requirements or a description of those conditions that must prevail before one could say the need has been met or the problem solved.

The needs-survey findings should be reported in practical terms, describing precisely the needs, the solution requirements, and the number and kinds of students involved. If the needs are not sufficient to justify further action, this report should be candid in admitting this. If this report calls for further action, the next step will be up to the principal. Later, this report should play an important role in selling the new program to various publics.

> Based on the needs survey, a plan for an experiential learning program, which presumably will require new forms of assessment, should be devised.

This initial plan, developed by the principal in consultation with the advisory groups, should describe the general outline of the proposed new program, including its philosophy and general objectives, and its relationship to the basic educational objectives of the school. As mentioned above, these programs should be directed toward learning experiences which are in keeping with the general objectives of the high school. (Clearly, not all experiences are learning experiences.) Experiences in the work-study or action-learning categories should strike a balance between participation and observation. Nontraditional learning experiences should be linked to formal school learning through parallel study programs within the school, counseling sessions, seminars, or formal contacts with a faculty adviser. Except in unusual cases, these learning experiences should be planned in advance and carried out under the direction of a faculty adviser. The usual practice is to formulate a contract between the school's program director and the individual student. The student's adviser monitors the student's progress in fulfilling this contract.

> With the initial plan drafted, a survey of community resources should then be conducted to identify opportunities for experiential learning in local businesses, facilities, programs, and agencies.

The final plan should be based on both the needs of the students and the learning opportunities available within the local community. Local resources should be identified in terms of learning opportunities available, interest in cooperative projects with the school, qualifications of personnel to work with students, and special problems such as transportation and insurance.

Examples of the kinds of local resources that should be investigated by school staff include other educational institutions, businesses with staff training programs, Junior Achievement, religious, fraternal, and social

organizations with training programs, cultural facilities such as museums or libraries, hospitals, scientific institutions, and governmental agencies at all levels. The best method of identifying and evaluating such resources is likely to vary from community to community. The principal and other school staff will obviously be important sources of information. An informal telephone survey of community leaders should also be useful. Any suggestions that are received will need to be followed up by direct contact with the group or institution that was thought to have the potential to provide some high school students with alternative learning opportunities.

> A review of plans with key school and community personnel should take place before final policies and procedures are set.

Before final and detailed policies and procedures are worked out and reviewed by the advisory group, the plans prepared to date should be presented for review and comment to the school committee, the superintendent, the faculty, student and parent groups, and to representatives of prospective participating organizations or agencies. This review should take place when the plans are sufficiently developed to permit useful discussion and approval yet are still open to changes which reflect the concerns and suggestions of the above-mentioned groups. Once the support of these groups has been obtained, a public information program that will explain the new program to the local community can begin.

> The design and supervision of the evaluation and crediting of experiential learning should be the responsibility of the School Evaluation Committee.

Because evaluation plays such a central role in experiential learning programs and because of the major evaluation issues created by these programs, the school, which has the responsibility for evaluation and awarding of credit, should form a special committee to deal explicitly with the evaluation component.

Early in the planning process, the principal should organize a School Evaluation Committee (SEC) consisting of the principal and heads of the departments and special services (such as library and guidance) that will presumably be involved in nonformal learning programs. The SEC will have the responsibility for determining the procedures and guidelines for the selection or development of all tests or assessment instruments and will determine the requirements for the awarding of credit. The SEC will select the appropriate assessment procedures for the various nonformal learning programs, supervise the execution of these procedures, and stipulate what learning is required for evaluation purposes.

At the outset, the SEC will draft its plans for the evaluation and crediting of nonformal learning, which will be reviewed by the faculty and by evaluation specialists in the school district office. These plans might also be circulated to other schools working on nonformal learning and assessment. (This same committee should be familiar with the evaluation and crediting procedures followed in the school's traditional learning programs. All too often new programs are subjected to demands that are seldom applied to existing long-standing programs.)

In each subject area involved in the nontraditional program a two- or three-member area committee might be set up with the department head as chairman. Following the guidelines and specifications established by the School Evaluation Committee, these area committees will be responsible for selecting or developing appropriate tests or other assessment procedures in their own subject areas and formulating the performance standards for awarding credits or honors in their own subject areas. (This Handbook will provide detailed directions to these test committees for this technically complex task.) The work of these committees will then be reviewed by the members of the SEC, who should be assisted by a measurement specialist. In effect, these area committees serve as subcommittees to the SEC. If it is necessary to obtain the approval of the tests by an outside agency, the principal will be responsible for submitting these tests and negotiating their approval.

Policies and procedures for the nonformal learning program should be developed in certain key areas.

The principal or the director of the nontraditional program appointed by the principal should formulate--in consultation with the advisory group and, where appropriate, the SEC--policies and procedures in the following areas:

Student Eligibility

The eligibility of students might be based on motivation, self-selection by students, personal interests and skills, level of maturity, or recommendation of faculty adviser and parents. Where space is limited, students might be selected on the basis of estimates of the value of the program to them. Selection should be made by a committee consisting of a faculty member, program director, and a student who is successfully participating in the program. Previous academic performances may or may not be a useful predictor of success in a program. Procedures and deadlines for student applications should be clearly stated. Students, for example, might apply for work-study programs through their counselor which a signed release from all of the other teachers involved. A pre-program orientation or counseling session might be required of all students.

Program Requirements and the Contract

Program requirements are usually set forth in terms of a contract between the school staff and the student. Attendance and conduct standards should be given attention along with reporting (logs, seminar reports, and so on) and evaluation procedures. The contract should describe the objectives and expected outcomes in ways that guide but do not unduly limit the program. It may include a statement of the student's purpose in undertaking the experience to serve as a reminder and guide as the student proceeds through the experience.

Both the student's special strengths and weaknesses may be stated. These may help focus the experience so that the greatest value is derived from it.

A list of alternative means for achieving the students' goals may help them make mid-contract adjustments with ease if the expected outcomes become jeopardized. Also, alternatives that vary in the kinds of media used may attract the students to richer experiences than they had anticipated. The duration and intensity of the experience should be set forth so that students are not unrealistic in undertaking a project that may require more time than is reasonable. Books, films, and other materials related to goals may be included in the contract so that students have the best opportunity to develop a significant knowledge base.

The adviser's role, the number and details of adviser/student meetings, and the purpose of the meetings should be specified. How the evaluation will proceed should be delineated; what is to be accomplished, how it will be observed, the criteria to be met, and the schedule for evaluation sessions should be described.

Staffing

Community-based educational programs and alternate modes of assessment and credentialing require faculty orientation, changes in faculty schedules, and increased student access to counselors. Thus, an increase in alternative educational programs would not necessarily diminish the number of staff but would involve expanded roles and, perhaps, changing functions of teacher-advisers and counselors. The success and quality control of nonformal learning programs depend upon the extent and quality of counseling as well as the extent to which the school provides for the coordination and evaluation of the new learning approaches.

In off-campus programs, the schools will, at times, be depending on important evaluation judgments from persons who are not certified teachers. Steps should be taken to determine the qualifications of these individuals by letters of recommendation, a visit by a local team of resource persons, or other means. For example, for the sake of those students involved in a work-study program in auto mechanics or cosmetology, the schools should have on record evidence of the qualifications of the mechanics and beauticians involved in the program. The school should also provide an orientation program for these members of the extended faculty so they can be certain just what the work-study program requires of them. The work of this extended faculty should be reviewed regularly by competent resource persons selected by the school. Presumably, these resource persons will serve in a volunteer capacity, if the demands on their time are kept to a minimum.

Any cost-benefit review of these out-of-school programs should take into account the value of the services performed by students in community-based programs, which may offset staff costs in whole or in part. Providing essential services, repairing or constructing public facilities, or developing new programs based on student services provide significant rewards to the community.

Transportation and Insurance

Questions about transportation and insurance are special concerns for action-learning, work-study, and other off-campus educational programs. While schools may be reimbursed for transporting students to and from school on a daily basis under the formula provided in Section 7A of Chapter 71 of the General Laws, other activities such as field trips and work-study programs

are excluded from Section 7A. However, field trips and other transportation costs may be included as part of the reimbursable expenditures under Chapter 70 of the General Laws.

With respect to insurance, schools should require students to enroll in the school insurance program if they intend to participate in off-campus programs, assuming that such insurance is available.

Reporting and Record Keeping

Clear procedures should be devised for reporting and recording non-traditional learning experiences on school and student records. In some cases, for example, community service or action-learning programs are translated into social studies credits and recorded accordingly; in others, they are simply recorded as credit hours for "Community Service." The student's faculty adviser might, for example, be assigned the responsibility of determining and recording a grade for the student's performance in a nonformal learning program.

One Valuable Outcome

One foreseeable and valuable outcome of launching a new program in nonformal schooling and its assessment will take place with the advisory group and the School Evaluation Committee. Pressed by the demands and questions stemming from these new approaches, these groups will be drawn into an extended study, discussion, and reflection on the nature and purposes of adolescent learning and the real goals of a high school today. In keeping with the warning of Socrates that "The unexamined life is not worth living," a fully functioning School Evaluation Committee and advisory group will make certain that the high school is not living an unexamined life.

REFERENCES

1. Woods, T. E. The administration of educational innovations. Eugene, Oregon: University of Oregon, 1967, 16.

2. Goodwin Watson (Ed.), Concepts for social change, Washington: NEA-NTL, 1967.

Other References

Berte, N.R. (Ed.) Individualizing education by learning contracts. New Directions for Higher Education, 1975, 3 (No. 2).

Eriksen, A. & Gartz, J. Partnership in urban education: an alternative school. Midland, Michigan: Pendell Publishing Co., 1974, 166.

Heussenstamm, F. K. Human services education: promises and problems. NASSP Bulletin, 59, No. 387, January 1975, 55-63.

Oregon Board of Education. Oregon graduation requirements: guidelines for record keeping procedures and sample forms. Salem, Oregon: Oregon Board of Education.

CHAPTER 5

SETTING PERFORMANCE STANDARDS

Overview

One difficult task facing schools as they seek to evaluate the experiential learning of students is that of setting performance standards. It is essential not only to decide what assessment method is to be used but to follow some systematic and carefully specified procedure to set a performance standard.

Even though schools have long experience with the processes of evaluating students and of awarding credit, the task of setting performance standards has not been studied closely by either teaching or administrative staff. Usually an individual teacher serves as the sole or the primary judge of a student's performance in the traditional high school program. Evaluation procedures for experiential learning programs are likely to face far more critical review than do those used in traditional programs. Parents and other community members seem inclined to question experiential learning programs on the grounds that the school may not be exercising an appropriate monitoring function. If schools are to continue offering alternative learning experiences that meet students' needs, the problem of setting meaningful and defensible performance standards with the cooperation of the community will require thoughtful attention.

This chapter raises some issues that must be dealt with when setting performance standards and reviews a number of possible methods that can be employed.

What Is a Performance Standard?

A performance standard is the minimum acceptable level of performance that will result in a student receiving credit. Traditionally the standard is equivalent to a "passing score." When the assessment procedure is an examination, the extreme limits of student performance are obvious: Maximum performance is demonstrated by those who answer every test question correctly; minimum performance is demonstrated by those who answer every question incorrectly. Difficulties arise in allowing departures from maximum and minimum performance when establishing performance standards, but common sense argues that such departures are necessary in the real world. The problem, then, is deciding where to draw the line.

As with many problems in measurement, a variety of attempts at solutions are available. There is no one ideal method of setting standards under all conditions. All possible procedures depend on subjective judgment at some stage in their application.

In the classroom, the teacher has responsibility for deciding what constitutes a "passing" performance. Sometimes the teacher announces a standard

without clearly indicating the basis for this judgment. The more common situation, though, is for the teacher to specify that classroom participation, projects, reports, test scores, etc. will each make a certain contribution to a total of points for each student. The final decision regarding the minimal number of points required for a passing grade typically is a normative one. The failing students, if any, will be drawn from the bottom end of the distribution of points. Variations in assignments or tests, however, usually require that a new passing level be set for each class.

The experiential learning situation often fails to provide the same kind of normative basis for setting performance standards as does the more traditional classroom program. The teacher-student learning contract or out-of-school work/community service program is usually tailored to a particular student and set of learning opportunities. The student must be evaluated on a more absolute basis than simply determining if he or she has learned as much as other students. Some combination of process (what the student did) and outcome (what the student learned) is likely to enter into the judgment. In the experiential learning setting, there will continue to be a normative basis for standards, since any reasonable set of expectations for student performance will be based on knowledge of what most students are capable of learning. In many instances, though, the normative base is in the minds of judges who set standards for a piece of work based on their experience with students who have carried out such work.

Some Properties of Good Performance Standards

When planning and carrying out procedures for setting performance standards, three related goals to keep in mind are fairness, realism, and consistency. Each of these goals will be considered in turn.

Turning first to the question of fairness to the student, it is important to develop procedures that permit adequate quality control of alternative learning experiences but that do not lead to high rates of failure for students who have worked conscientiously over a period of time. In this connection, some attention should be given to the consequences for a student who fails to meet a performance standard under various types of learning programs and assessment procedures. These consequences can range from minor inconvenience to major disruption. At the minor inconvenience end might be the case where a student challenges an existing course, receives permission to take an examination, and fails to achieve the performance standard. If the student has not invested a significant block of time preparing for the examination, his or her failure to meet the standard may be of little moment. The implications of failure, however, will be quite different when the student has followed a carefully planned program covering several months. If the student has engaged in a set of activities that were the results of a learning contract between him and a teacher/monitor, that student can reasonably expect to receive credit for this learning activity. When group projects are involved, attention needs to be given to the relative contribution of each group member. If the role of each student has been clearly specified, the standards-setting procedure can be carried out to determine the extent to which the student has fulfilled this role.

If credit is not to be an automatic consequence of any alternative learning experience, what steps can be taken both to maintain quality standards and still make failure a rare occurrence? Some form of ongoing evaluation will perhaps best serve this need. When programs are being designed and learning

contracts prepared, the teacher and student should agree upon checkpoints at which the student would report to the teacher the progress that he or she has made. This will enable the teacher to detect potential problems in time for the student to revise his approach so that it will be possible to award credit when the alternative learning experience has been completed. If the student is unable to change his or her behavior and continues to fail to meet the requirements of the learning contract, it would then be both fair and reasonable to deny credit for the alternative learning experience.

Ongoing monitoring of an alternative learning experience will be of great value to the student. However, a fair evaluation system also needs to be based on a realistic view of what a student can learn in an experiential program. A common mistake is to establish standards for students that are unrealistic. This is most likely to happen when the individuals setting the standards do not have direct knowledge of the performance of other students on the type of task under consideration. What the school says students must attain to receive credit should be clearly related to what a typical group of students performing at a reasonable level of effort should be expected to achieve. Since school and community people should work together in planning and developing experiential programs, all parties concerned will need to review evidence of what the average student can learn in each program offered. Interviews with students and reviews of student papers or products are helpful ways of getting a fix on what is realistic to expect of students. Some attention will need to be given to an evaluation of the extent to which students are applying themselves in any given experiential learning setting.

Meeting the joint requirements of fairness and realism will be a challenging job, but it is also essential to achieve a high degree of consistency in the implementation of standards. The greater the number of students involved in experiential learning programs, the harder it is to achieve and maintain consistency. It is clear, though, that growth in size of such programs means that more students' lives are affected, so consistency becomes more important as a consequence.

What can a school through its evaluation committee do to monitor the consistency with which performance standards are applied? Some monitoring procedure is clearly needed, such as a periodic review of the results of applying a particular set of assessment procedures and standards. It may be useful to have a member of the school evaluation committee sit in on some assessment sessions. An initial set of standards may have been applied consistently but may be either too high or too low. Alternatively, standards may have drifted over time to become either higher or lower than was originally intended. The rate of failure within a particular program should be observed closely for changes of this nature. Such changes may be quite appropriate; a review should merely establish that the changes are consistent with the objectives of the alternative learner-certification system. It may be the case, however, that the school and community staff who are evaluating a particular experiential program need to review and revise their procedures. Merely reviewing failure rates is not a sufficient monitoring procedure since standards might be relaxed over time in order to maintain a low failure rate. The School Evaluation Committee should receive reports routinely on the basis of which it can make judgments about the adequacy of the evaluation process.

Only by careful initial planning and conscientious follow-through and

follow-up can a school meet goals of fairness, realism, and consistency for performance standards for experiential learning programs.

Methods of Setting Performance Standards

In the following pages, six methods of setting performance standards are described and evaluated:

Traditional or Arbitrary Method
Review of Test Questions or Tasks
Identification of Prerequisite Skills
Use of Normative Data
Relationship to Course Performance
Use of Preselected Groups

Each of these methods is described and evaluated in turn.

Traditional or Arbitrary Method: When dealing with an examination or set of performance tasks, the worst of all possible methods is to decide on an arbitrary percentage of correct answers because that percentage has traditionally been associated with a "passing" score. The method has its roots in the percent grades used almost universally in schools and colleges until about 1920. Unfortunately, this approach (the "judgment uncontaminated by data" method) is the most frequently used process for setting performance standards. In this approach, one attempts to determine what seems right for a passing score. The answer usually represents a compromise between those who desire as high a percentage as possible to show their interest in "quality" education and those who are afraid that their standards will be impossible for the bulk of students to meet.

Clearly, any method that fails to take into account the nature of the learning experience, test questions or tasks, and students is unsatisfactory. A performance standard such as 65 or 80 percent correct might be reasonable for certain tests for certain groups of students; it may be wildly inappropriate for other tests with other students. As a rule of thumb, blanket performance standards should be avoided.

Review of Test Questions or Tasks: This method sets a performance standard for each set of questions or performance tasks used for a specific learning experience. Informed judges examine each question or task and decide if a minimally competent student would answer the question correctly or perform the task adequately. The performance standard is simply the average number of questions or tasks that each judge has determined should be answered correctly by minimally competent students.

An improvement on this procedure is to ask each judge to review the records of a number of minimally competent persons and estimate the proportions of such students who would answer each question correctly. The sum of these proportions represents the minimally acceptable score. One of the advantages of this procedure (estimating the number of students who answer each question correctly rather than identifying the particular questions that could be answered correctly) is that the effect of guessing can be acacounted for. In a multiple-choice test, the number of choices affects the difficulty of each question and the likelihood that students will guess some

answers without knowing the concept or skill being tested. Judges can adjust for this by estimating higher proportions correct for questions with few options and lower proportions correct for questions with many options.

The method of estimating proportions of a minimally competent group who would give correct answers takes into account the possibility that individual students will miss some questions that are considered fairly easy but will give correct answers to harder questions. All questions are given some weight in setting the performance standard.

As an example, consider a hypothetical test of 10 questions which were evaluated by one judge as follows:

Question Number	Estimated Proportion of Group of Minimally Competent Students Who Would Answer Correctly
1	.80
2	.85
3	.75
4	.70
5	.65
6	.40
7	.50
8	.45
9	.40
10	.50
Sum of Proportions =	6.00
Performance Standard =	6 or more questions correct

The hypothetical example involves one judge whereas three to five are recommended. The performance standard is obtained by averaging the results from all judges.

When dealing with student products or performance skills rather than test questions, judges will need to decide on the relative weight to be given to the different aspects of the product or performance (See Appendix H.)

Identification of Prerequisite Skills: This method is of most use when attaining an ultimate goal is unlikely without having acquired skills. Success at long division, for example, is not likely unless the student can subtract and multiply. It should be recognized that most courses or experiential learning programs do not have the hierarchical progression of mathematics. The identification-of-prerequisite-skills method, though, is most appropriate when decisions can be made about a student's ability to undertake a new unit in a somewhat hierarchical sequence of units.

This method uses pooled judgments, but the decisions are made on the basis of whether or not success on a question or task requires knowledge or skills that are essential to the future application of skills learned in the course or program. The method, in practice, has the same advantages and problems as reviewing questions or tasks. But the focus is not on asking if a minimally competent person would get a particular question correct, but rather if success on a question is required before some specified future activity is possible.

A task analysis of the future activity can help to determine what behaviors are required in the performance of complicated tasks. (See references 1 and 2 for a discussion of applications of the task-analysis approach to performance evaluation.)

Use of Normative Data: A standardized test or a performance-assessment procedure can be administered to students and the results used to set a performance standard. Even though this method has empirical elements, pooled judgment is still needed both in the selection of a test and in the selection of an appropriate norming group. A norming group must be chosen that is sufficiently similar to the group in question. In contrast to previously discussed methods in which judgments concerned questions or tasks, judgments in this method concern which students in a class or group are at a minimally competent level. The test or performance-assessment procedure is administered to a group of such students. The mean score or average performance of that group may be taken as the minimum performance standard. Using normative data tends to insure the setting of a performance standard that is realistic in terms of the percentage of the population that is judged to have reached an adequate level of performance on the basis of the test. This method is most appropriate for situations in which an objectively scored examination is the basis for awarding credit.

Relationship to Course Performance: This method of setting a performance standard is most appropriate for making decisions about credit when an alternative learning experience is related to the content of a particular course. A test that has been selected or developed as the basis for awarding credit is administered to students completing the regular course. Scores on the test are then related to grades earned in the regular course. A table or graph can be prepared showing the relationship between test scores or performance ratings and grades earned in the regular course. The following is an example of such a table. It is based on the spring 1974 administration of the Social Sciences section of the Sequential Tests of Educational Progress to a group of Massachusetts seniors:

Grade	Mean Test Score	Number of Students Earning Each Grade*	Percentage of Students Earning Each Grade
A	62.1	212	16.4
B+	57.2	210	16.2
B	52.5	297	23.0
C+	49.3	135	10.4
C/D+	44.0	334	25.8
D/F	41.7	106	8.2

*A total of 1,451 students were tested. No grades were reported for 157 of these students.

The data in this example were collected from a number of different schools whereas an individual school may have only data from its students to work with. School-wide data can be organized in a table which takes the form of a grid containing a number of cells. Inside each cell is a number based on the number of people who had scores on the test and grades in the course corresponding to the position of that cell. Usually the data are organized to allow direct readings of probabilities of earning a particular course grade given a particular test score. Readers with statistical backgrounds will recognize the fact that scores on the examination can also be related to course grades by means of a regression equation, which will give a predicted grade for any given test score.

After the test data have been summarized in the manner described above, the identification of the test performance standard is made on the probability of obtaining various grades. (For those who may wish further information, Willingham [3], in his book College Placement and Exemption, reviews possible methods of using course performance as a basis for setting a performance standard.)

Use of Preselected Groups: This method requires that information that is independent of the particular assessment procedure being applied be used to select one group of students who are passing a course and another group who are at a failure level. Instructors' judgments should be used to divide students into the two groups. In making these judgments, it is essential to differentiate course-related competencies from disciplinary considerations and other factors that contribute to grades.

The assessment procedure is then applied to both groups and the distribution of scores obtained. A performance standard is then selected that best discriminates between the two groups.

This method has the important advantage of inherent concurrent validation: The test is validated against independent informed judgments of student competency. Furthermore, this procedure allows the use of classical measures of test quality such as reliability, validity, item discriminations, and the like.

Error of Measurement

The methods of setting performance standards that are described in this chapter combine human judgment and empirical data in a systematic process. Regardless of the quality of the procedure used, however, any system will result in some error of measurement. This "error" is not attributable to arithmetical mistakes that could be avoided by more careful calculations. It is related to the complexity and variability of people and the difficulty of measuring their skills and knowledge. Even if you were somehow able to achieve complete knowledge of each student's competencies and had a perfect standards-setting procedure, you would still be able to detect some degree of error in any operational system.

There are two major types of errors possible in setting standards. If the standards are set too high, the probability increases that those who should receive credit will be classified failures. If the standards are set too low, the probability increases that those who should fail will be passed.

Clearly, the setting of standards must take into account the costs or losses associated with misclassifications of each type. In every situation, the basic principles for standards of fairness, realism, and consistency will apply, but it will also be necessary to consider the social implications of decisions about students' competencies. Consider the area of driver education, for example, where an incorrect judgment of a student's competencies might result in the licensing of an incompetent driver.

Final Thought

The job of setting appropriate performance standards is as critical as it is difficult. Members of the School Evaluation Committee should allow sufficient time in their projects for a thoughtful analysis of possible approaches. If at all possible, more than one method of setting standards should then be tried and the results compared and evaluated. A decision can be made after this evaluation regarding the method or methods that will be employed for each assessment procedure. Schools may wish to use outside consultants to assist in the evaluation process.

REFERENCES

1. Thornton, R.F. & Wasdyke, R.G. A taxonomy for the development of multidimensional test specifications. Princeton, New Jersey: Educational Testing Service, 1972.

2. Thornton, R.F. & Wasdyke, R.G. The development of criterion-referenced tests using the behavioral-consistency model. Princeton, New Jersey: Educational Testing Service, 1972.

3. Willingham, W.W. College placement and exemption. New York: College Entrance Examination Board, 1974. 132-199.

PART II: ASSESSING STUDENT COMPETENCE

CHAPTER 6

OVERVIEW OF ASSESSMENT METHODS

The earlier chapters of this Handbook have reviewed the context within which nontraditional experiential learning programs are now being carried out within the Commonwealth of Massachusetts. Attention has been given to the fact that existing educational legislation and regulations place primary responsibility upon school staff for the evaluation of student learning within such programs. Strategies have been identified for combining the professional expertise of school staff with the contributions of parents and other community members.

This chapter presents an overview of five assessment methods that can be used either separately or in combination to assess student learning. The procedures are designed to be used in evaluating student competencies for awarding high school credit. It is assumed that credit would be awarded only for planned educational programs sponsored by the school, even though much of the actual learning may take place outside the school setting. It should be recognized, though, that many of the principles and procedures described could be applied to other situations. In the traditional school learning programs, for example, it may prove useful to incorporate elements of one or more of these procedures when awarding grades in regular classes. Similarly, schools could draw on these methods for use in adult education programs. Some schools might also wish to offer the student the opportunity of obtaining credit for experiences that were not carried out under the school's sponsorship. Students who have entered a school from another region, state, or country might well argue that their prior experiences should make them candidates for some high school credit. Decisions about using any of these assessment procedures would, of course, be the responsibility of the School Advisory Group and School Evaluation Committee.

Summary of Assessment Methods*

The five assessment methods described and discussed in this Handbook, each in a separate chapter, include the following:

Testimony: The testimony method requires that a third-party observer such as an employer, coach, or employment supervisor certify or attest to the fact that a student has completed a predefined set of activities.

Committee Review: The committee-review procedure involves three or more judges who evaluate student learning outcomes that are not readily measured by paper-and-pencil tests and not easily observable.

Performance Assessment: The performance-assessment method refers to a process in which the student is asked to carry out clearly defined tasks under a carefully specified set of conditions.

---

* Figure 1 in Chapter 3 suggests assessment models for a variety of learning activities.

Product Assessment: Product assessment requires judges to evaluate a variety of student products such as drawings, paintings, musical compositions, science projects, diaries or reports of travels, photographic essays, or handcrafted items such as cabinets or bookcases.

Objective and Essay Examinations: The objective and essay examination assessment model employs a formal set of examinations and grading procedures to assess student competencies.

Reliability and Validity of Assessment Procedures

Each of the following five chapters on assessment methods describes development and review procedures that would permit adequate quality control of experiential learning programs and at the same time would be fair to the students whose learning is to be evaluated. Two technical requirements--reliability and validity--must be satisfied before any assessment procedure can be judged adequate for the student, the school, or the community. Since these characteristics apply to each of the assessment methods, they are discussed in this chapter and suggestions are made for further reading for individuals who will be closely involved in the job of designing and developing programs for evaluating student competencies.

Reliability: Reliability can be defined as an indicator of the dependability or consistency of any measurement or assessment procedure. If we were to assess the same set of people again and again with the same or comparable instruments or procedures, we would not obtain identical results on each assessment. This is because assessment, like any other human endeavor, involves a certain amount of error. The errors can be of two types: There is systematic error, such as when scores on an instrument are all biased in one direction--for example, a ruler that is too short, a scale that adds five pounds to each weighing, a test that counts a wrong answer correct. Scores on such defective instruments may be wrong, but they will be wrong consistently, dependably, predictably, and reliably. However, random errors, as the name implies, occur by chance. They are temporary and shifting, often due to unknown causes such as fluctuations in mood, health, motivation and so on. The greater the amount of random error in a measuring instrument, the less the reliability of that instrument. A good example would be that of two watches--one very fine and expensive, the other one rather cheap and poorly made. The fine watch will measure time reliably and dependably. Even if it is set to the wrong time so that it does not give an accurate estimate of the correct time, it will still be consistent, or, in this sense, "reliable." The cheap watch, on the other hand, will have a certain amount of purely random error. Sometimes it will be too fast, sometimes it will be too slow, and it may be impossible to predict which way the error is directed. It is this kind of random, unpredictable error that makes the cheap watch unreliable.

Reliability is a proportion, the proportion of the score that is a true measurement of the thing being measured. Obviously, the more random error, the smaller the proportion of truth in a measurement. The less truth in the measurement (the smaller the proportion of truth to error) the lower the reliability.

Unfortunately, there is no direct way of determining an instrument's reliability. There are, however, various methods of arriving at estimates of reliability. Most of these methods involve comparisons of two scores for all of the individuals in some defined group. The relationship between the two sets of scores is expressed in terms of correlation coefficients. A correlation coefficient is a number that varies between -1 and +1 and that indicates the degree of relationship between variables. If the correlation coefficient is +1, there is perfect agreement between the two variables. If one goes up, the other goes up a proportional amount; if one goes down, the other goes down a proportional amount. If the correlation coefficient is -1, the relationship is still perfect except reversed. If one variable goes up, the other goes down by a proportional amount. A correlation coefficient of 0 (zero) indicates no relationship between two variables.

Some suggestions are made in the following chapters regarding ways that schools can assess the reliability of the assessment procedures that they choose to develop and use. Unless schools have access to technically trained measurement staff, they will have to seek the assistance of individuals from the State Department of Education, local colleges, regional consortia, or the like in order to plan and carry out reliability studies of the assessment procedures that are used. See the reference list at the end of this chapter for titles of some basic texts on test reliability.

Validity: In addition to requiring that each assessment demonstrate its reliability or consistency, the validity of the procedure should also be established. The validity of an assessment procedure is the extent to which it measures what it is supposed to measure. The job of establishing the validity of an assessment procedure is much more difficult in the world of education when student learning is to be assessed than it is in the area of physical measurement where such properties as length and volume are to be assessed. It is fairly easy to get people to agree that a ruler is an appropriate measure of length; it often proves quite difficult to get individuals to agree as to precisely what it is that a test or assessment procedure actually measures. It is important, however, for a school district to undertake the task of establishing the validity of its assessment procedures and of adapting or even discontinuing any assessment procedure that cannot demonstrate its validity for the purpose for which it is used.

How then can a school district undertake to assess the validity of any assessment approach? Psychometricians have identified a wide variety of methods that could be used to assess the validity of an assessment procedure. The references for this chapter include several that discuss this topic in some detail. In the school setting, however, it should be sufficient to investigate content validity and criterion-related validity.

An assessment procedure can be considered to have adequate content validity if knowledgeable and experienced judges agree that it constitutes a reasonable sample of the kinds of skills and knowledge that were identified as the expected outcomes of a nontraditional or experiential learning program. The judges need to become familiar with the anticipated learning outcomes of the program and then to evaluate each of the questions in a test, each of the performance tasks in a performance-assessment setting, or the analogous elements of one of the

other assessment procedures. Since assessment procedures typically constitute only a sample of all the possible questions, tasks, or products that might be required of students, the judges need to decide whether or not an adequate balance is represented in the sample required of a particular student. Frequently, a systematic review of the components of any assessment procedure will reveal that some of the included elements are indeed not essential whereas certain significant areas of student learning outcomes have been overlooked.

In order to establish the criterion-related validity of an assessment procedure, one must compare the decision made about a student using this assessment procedure with an independent decision about that student made from different information. If a school is using a performance-assessment approach to decide whether or not a student should receive credit for an experiential learning program that required him or her to master a certain set of skills, it needs to seek other information as to whether or not the student has indeed mastered those skills. A three-judge group might have asked a student to carry out three or four specific tasks under carefully controlled conditions and on the basis of this information, either decided to award or not award credit. This type of performance assessment must take place within a certain period of time on a certain day. If it proved possible to observe the student performing these precise tasks in a real-life situation over a period of several days or weeks, it would very likely be possible to state with some degree of certainty whether or not the student really had mastered these skills. It would then be possible to compare the decision reached in the school's assessment procedure with the decision that would have been reached if the more extensive body of information had been obtained over several days or weeks of observation.

The task of establishing the criterion-related validity of an assessment procedure will be most practical and efficient when the school can use independent data already available. When considering the use of an objective examination that is related to the content of a particular high school course, it will be possible to administer this test to students who have taken the course and to compare scores on the test with grades earned by the students. It should be recognized that this method of establishing the criterion-related validity of a test can be carried out at the same time information is collected on performance standards for the test. The discussion in Chapter 5 on the use of normative data in the establishment of performance standards points out that a test would have to be administered to students who had taken an appropriate course.

For more information on validation, see the discussion by Lee J. Cronbach in Validation of Educational Measures.

Implications for School Staff

As experiential learning programs grow, they will make heavier demands on teachers' skills as counselors and monitors of their students. Thus, if experiential programs become a major part of a high school's offerings, it will be essential to take into account the demands on teachers and other administrative staff who are associated with such programs. Growth in experiential programs will make it essential to adjust the teaching load of teachers to take into account their contribution to program development and to the monitoring and evaluation of student learning.

In addition to considering the total amount of time that will be required by experiential learning programs, it will also be necessary to schedule the assessment activities so that they can be integrated into the other instructional and administrative demands on staff. It will clearly be undesirable to attempt to carry out all assessment procedures at the end of the school semester or year. Instead, a program of assessment should be scheduled so as to minimize conflict with other major demands on teacher and student time. It may be useful to review early in the school year the total number of evaluations that will be needed under each assessment method. Decisions could then be made regarding the number of different assessment sessions of each type that will have to be scheduled. As with other school activities, it will be essential to announce deadlines early and to issue reminders to students, teachers, and cooperating community members.

REFERENCES

For some additional background on the concepts of test reliability, the following basic textbooks are recommended:

Cronbach, L.J. Essentials of psychological testing (3rd ed.). New York: Harper & Row, 1970.

Thorndike, R.L. & Hagen, E. Measurement and evaluation in psychology and education (3rd ed.). New York: Wiley, 1969.

Those interested in a more detailed overview of theoretical and statistical approaches to reliability should see:

Stanley, J.C. Reliability. In Robert L. Thorndike (Ed.) Educational Measurement (2nd ed.). Washington, D.C.: American Council on Education, 1971.

For more information on validity, see:

Cronbach, L.J. Validation of educational measures. In Robert L. Thorndike (Ed.) Educational Measurement (2nd ed.). Washington, D.C.: American Council on Education, 1971.

An excellent source of information about the evaluation of experiential learning is the Cooperative Assessment of Experiential Learning (CAEL) project conducted by Educational Testing Service, 10 task force colleges and universities, and a larger CAEL Assembly of member institutions. The purpose of the project is to develop and try out materials that will be useful in strengthening the assessment of experiential learning in higher education. CAEL has produced a number of guidebooks for student and faculty use that treat different facets of learning assessment. Some of these publications are referenced in later chapters of this Handbook. Schools can obtain further information about CAEL by writing to CAEL Assembly, Educational Testing Service, Princeton, N. J. 08540.

CHAPTER 7

TESTIMONY

The scene is a ski school. Students are packing their gear in preparation for their return home. A ski instructor is filling out a report form for a high school junior who has successfully completed the basic cross-country skiing course at the school. When the student presents this report to his faculty adviser, he will have fulfilled part of the requirements for credit in physical education.

Definition

Testimony may be considered both a method of assessment and a basic part of some of the models to be developed in later chapters. It involves both affirmation by a learner and, wherever possible, certification by a third party such as the ski instructor in the scene above that a plan of action agreed upon in advance with a faculty adviser has been carried out in the manner and under the conditions stated. This method could be employed in some situations where it is impossible or impractical to have the entire learning activity observed by a third party such as an employer, coach, or instructor.

Typical areas for use of this method with only partial third-party observation would be in the personal-development skills--cross-country skiing, backpacking, or career exploration. In each of these, the consequences of gaps in the evaluation system that permitted student laxity to go undetected would fall largely on the individual student. The same would not hold true for auto mechanics, driver education, or other areas in which individual failure undetected could endanger public safety.

A third party would not necessarily be required to evaluate the skill or effectiveness of a student's performance but only to certify that he or she has completed the tasks.

Description of the System in Operation

The school identifies those areas in which a student may earn credit through testimony by independent work and sets limits on the number of credits in this area. For example, up to one quarter of the physical education electives may be replaced by contracts for equivalent physical activity under the supervision of an acceptable adult. This person could be a scout leader, coach, dance instructor, camp director (for summer make-up work), or a parent. The nature and duration of the activity would be specified, along with the dates and nature of the certification. The student would be responsible for insuring that the plan was approved by the appropriate adviser prior to

the start of the semester and that certification forms were filed on schedule.

Possible Applications of the Testimony Approach

The testimony assessment approach will be most appropriate in those action-learning settings that involve participation in well-established community service or extracurricular activities. The Junior Achievement program or a hospital service program would be typical examples. In such settings, it should be possible for the faculty adviser to familiarize himself with the conditions under which learning will take place and to become acquainted with the adults who will have an opportunity to observe the student in the action-learning setting. Where well-established programs exist and there are easy opportunities for communication between faculty members and staff in the outside agency, the testimony approach can be used with confidence and with minimum demands on faculty time. When the student's work will be largely unsupervised or when there are only limited possibilities of obtaining information from other adults, a greater burden will fall on the faculty member and it may prove necessary to require the student to supplement his "hands on" experience with written reports or other evidence that learning has indeed taken place along the lines anticipated. The greater the difficulty that will face an individual faculty member who must evaluate a student, the more likely it will be that committee review, as described in the following chapter, should replace testimony as the assessment method.

For a description of action-learning programs in 25 schools, see the National Association of Secondary School Principals' report, Twenty-five Action-Learning Schools (1).

Strengths and Limitations

The testimony method will make heavy demands on the professional judgment of the individual faculty adviser. At the same time as it offers considerable opportunities to design flexible programs tailored to the needs of students and the available learning options in the community, it will require the development of policy guidelines that retain the faculty member's flexibility but provide for uniformity in the application of quality standards. Collaboration among the faculty members in a given department will be essential. Provision should be made for faculty planning and training sessions that permit the faculty advisers to reach agreement on the nature of the contracts and the level of monitoring that will be appropriate in different types of learning situations. In addition to reaching agreement on an initial set of procedures for evaluating students, faculty members should hold at least annual reviews of the procedures they have been following and try to recalibrate their internal standards so that as high a level of consistency across teachers as possible is maintained.

REFERENCE

1. Twenty-five action-learning schools. Reston, Virginia: National Association of Secondary School Principals, 1974.

CHAPTER 8

COMMITTEE REVIEW

The scene is a walnut-panelled office in the downtown headquarters of a large corporation. Manuel, an 18-year-old executive intern, is seated at his desk completing a preliminary review of the poverty indices published by various public agencies. The evaluation, which was requested by the Community Relations Department of the corporation, is a very challenging task, even for an exceptional mathematics student like Manuel. Elsewhere in the building, Kathleen, a 17-year-old intern with a public agency, has been helping redesign certain offices in order to conserve energy. Both of these high school students are maintaining logs of their experiences and have made copies of the work they have completed. Later in the year, each will submit a report with supporting materials for evaluation to a committee of high school staff and community members. Since they are both working conscientiously, are meeting the requirements of their learning contracts, and are keeping their faculty adviser informed, they will almost certainly earn credit for their internship experience.

Definition

The committee-review procedure is one in which several people evaluate learning outcomes that are not readily measured by paper-and-pencil tests and not easily observable. Examples of such outcomes are those from an executive internship program such as the one described above or a government internship program that has involved individuals in the daily routine of legislators or other government officials and under such conditions that observations of the student need to be carried out over a prolonged period of time. Otherwise, the observer is not confident of seeing a representative sample of the student's many and complex behaviors. Observation of the student's performance in such circumstances by school staff is not an efficient method of evaluation.

On the other hand, paper-and-pencil tests, which are efficient in measuring a large sample of complex behaviors quickly, may not be valid. The lack of validity could stem from the fact that most such tests are based on what students are taught about government

in a traditional classroom. Consequently, the questions are likely to cover vocabulary such as "checks and balances," the historical functions of government agencies, constitutional provisions, methods of apportioning representation, and so on. But in the internship, the concepts that are learned are more limited in number than those learned in a classroom but frequently learned in greater depth. To test the validity of an internship experience, a test might have to consist of 50 questions on a limited topic such as the development of a political campaign, whereas most classroom tests may have only one or two such questions. Since the latter tests may be invalid for evaluating internship experience and the former unavailable, the evaluation requires a different procedure.

This chapter reviews a number of issues relating to the staffing and operation of an assessment committee. It should be pointed out that the ideas presented are also applicable to assessment situations in which a committee focuses primarily upon observing a student's performance or evaluating specific products without relying extensively on supporting documents. It will be useful to consider the ideas in this chapter when reading Chapter 9 (Performance Assessment) and Chapter 10 (Product Assessment).

Description of the System in Operation

As with each of the other assessment methods recommended in this Handbook, policies covering elements such as provision of information, student eligibility, and application procedures will be developed by the School Evaluation Committee.

In a sense, the committee-review assessment method is like developing a unique test of each individual's learning experience. A committee examines the student's work or the student himself and measures the amount and kind of learning represented by materials or presented by the student as evidence.

The choice of committee size should be balanced between a group large enough to insure an adequate level of reliability of the committee's judgment and small enough to permit easy communication and decision making. A committee of three to five will serve most evaluation purposes. One of the members should be the student's faculty adviser. The community will often provide additional members. Every other member's term should be staggered so that no committee is ever in session without at least one experienced member in addition to the adviser. The member with the longest service may serve as chairman, or the chairman may be that member whose discipline is most clearly related to the student's experience.

Portfolios: The portfolio, narrowly defined, is a file of information about the student's experiences. It is the vehicle for organizing the experiences into manageable, assessible form. More broadly, the portfolio is a process by which experience is translated into educational outcomes. It consists of an agreement between the

student and the adviser about the student's goals (learnings), the method (experience) by which he will attain the goals, the timetable, the materials to be produced, and the reward (credit to be granted) for attaining the goals. Another section of the portfolio consists of summaries of discussions which the student's adviser has had with the student. A third section consists of the documents--writings, work records, licenses, press clippings--which the review committee must evaluate.

Two helpful resource documents for schools interested in the portfolio approach are A Guide for Assessing Prior Experience Through Portfolios (1) and A Student Handbook on Preparing a Portfolio for the Assessment of Prior Experiential Learning (2). These reports are focused on the prior learning of college applicants, rather than the school-directed learning of high school students, but they treat a number of issues common to both contexts.

The Committee-Review Process: The committee's job is to decide the extent and level of learning that is represented by the documents which the student presents for evaluation. The committee must decide whether the learning evidenced by the documents presented is worthy of credit. Since judgment is the means for making the evaluation, all efforts to eliminate bias should be made. Before the judgments are made, the committee should agree on the criteria to be applied. For example, if the student submits copies of lesson plans, anecdotal records, and supervisor ratings stemming from a child-care experience, the committee should decide the degree of skill or proficiency which should be indicated. The skill of a master teacher would be too high. The knowledge about teaching of the average high school student who has never studied teaching or psychology is probably too low. The committee must decide the point in between that is acceptable. It must also decide the components of the total experience that should be judged separately. In the child-care example, the committee may want to judge separately the evidence of the student's learning about child development, recreation facilities for preschool children, planning learning experiences, and interpersonal relationships.

It should not be assumed that students attend the committee review session, although they should be allowed to if they wish to supplement the documents with comments or if the members of the committee want to ask them questions. Their presence would be necessary in certain other instances as when documents are not available--for example, some component of a day-care experience, such as learning about theories of child development, during which students might have had no opportunity to obtain documents to give evidence of their knowledge. In other instances, a goal of the experience might have been to increase the students' speaking abilities. In both cases, the committee-review technique is that of an interview. However, different emphases are given in evaluating the students' response to the interview. In the former case, emphasis would be on the students' knowledge of the content of child-development theory. In the latter, emphasis would be on oral expression.

Development

Planners of committee-review procedures confront three basic problems: ambiguity, focus, and criteria. The first two of these are particular problems when conducting interviews. If questions are

ambiguous, students may respond to what they take to be the question rather than responding to the intended question. Developers should try out questions if possible, but in any case, should keep records on their thoughts about each question as it is used. In this way, a file of questions that work will eventually be developed.

Focus has to do with attending only to the specific behaviors to be assessed and not irrelevant ones such as appearance. Developers of review procedures should make explicit the types of questions that should not be asked. One such example might be "How did you like working in that political campaign?" The question would be a valid one (though ambiguous) only if the objective of the experience was to develop positive (or critical) attitudes toward political campaigns. Otherwise, the question should not be asked at all or only after the evaluation when the student's answer will not influence the committee's judgment.

Criteria have to do with the standards against which student achievement is to be mastered. The committee must agree on a reference point which each member has in mind as the evaluation is carried out. The committee may ask closed questions (agree/disagree or other alternatives) which are useful for classifying a student's attitude or perception, open-ended questions which permit the committee to probe for depth of knowledge, and scale items which allow students to express degrees of agreement or disagreement. Scale items are frequently used as checks against open-ended questions. Regardless of the type of question, all should follow the same rules. Questions should:

1) be clear and unambiguous.
2) be directly related to the skills and knowledge being assessed.
3) be appropriate for the kinds of information to be obtained.

Questions should not:

1) demand knowledge or information the candidate does not have or should not be expected to have.
2) include leading questions.
3) require personal information that the candidate may not wish to reveal.
4) be slanted so that the candidate offers the socially acceptable response.

When information is not a matter of record or easily obtainable, and when there are many students whose experiential learnings must be evaluated, time-saving procedures will have to be developed. One of these is the use of ratings. Ratings are a way of quantifying opinion and, as such, lend themselves more easily to computations than do verbal assessments. A checklist, one form of ratings, helps the committee reviewers focus on particular descriptions of behavior which are placed on a particular point on a scale. In developing rating scales, the committee will have to decide the behaviors that are likely to overlap the various experiences that students may choose. For example, it may be decided that all out-of-school experiences should lead to more knowledge and understanding of adult roles. This, then, would become one of the scale items to be developed.

One rating form is composed of between 5 and 15 such scales and is suitable for many purposes. Each scale point should have a specific behavior description attached to it. A five- or seven-point scale usually gives enough discrimination, and too many levels confuse raters. In summing the scores on individual scales, the scales measuring the most important factors should be given the most weight.

The task of developing committee-review procedures requires work at the departmental level following initial guidelines set down by a school evaluation committee. It will be necessary to deal with a number of practical and theoretical issues. For example, there will often be a temptation to broaden the alternative educational experiences that are open to students even when some of those experiences do not lend themselves to ready evaluation. Moreover, teachers tend to require more evidence of learning from students involved in special programs than they do from students in regular classrooms. Consistency of standards is indeed important, but fairness and realism of standards need also be assured.

Strengths and Limitations

The most important advantage of the committee-review technique is its flexibility. Once procedures are established, and particularly when questions are developed, tried out, and improved, the committee is in a position to move from the evaluation of one type of experiential learning to another. Naturally, the wording of questions must be adapted to the specific situation, but much of the development effort will not have to be duplicated. The committee-review technique, particularly when it uses the interview approach, makes possible a probing for in-depth information that other evaluation methods do not permit.

The major disadvantage of committee review is its demand on faculty time. Whereas a test, once developed, only involves staff in a minimal way, evaluation by review committee involves staff heavily. This will be true even when there is a good deal of community participation in the committee-review process. The review method is not recommended when many students must be evaluated.

Technical Considerations

As noted in a previous chapter, reliability is a key issue in any measurement activity. The dimensions of even physical objects when measured by different people with the same yardstick might appear to be different. Yet we know the calibrations of the yardstick do not change with the users. However, the scale points on a rating instrument do not mean exactly the same thing to different people. One rater may give a student a high grade on the first factor and an average grade on the second, while another might give an average grade to the first and a high grade to the second. Which is the true measure? It is probably closer to the averages of the two raters than to either one alone. For this reason, it is essential that the review of portfolios or other evidence or learning or the interviewing of students be conducted by more than one judge and preferably by at least three.

It may be argued that teachers almost always decide alone what grades their students' work deserves. The difference between the classroom situation and the experiential learning situation are important. In the first place, teachers give their students many grades -- for quizzes, end-of-term tests, and homework and class recitation. Any bias the teacher may have at one moment is probably balanced out by other biases at other moments. In the second place, teachers grade a student's work in the context of work submitted by others. It is much easier to put many pieces of work in order according to quality than it is to decide whether one piece of work is good or bad.

There is no doubt that the committee review procedure is demanding of staff time. However, the time is required in a concentrated period. It is possible that the total amount of staff time given to the evaluation of a student's experiential learning outcome is less than the time that would have to be given to the same student in a classroom situation. In any event, the task of reviewers will be made easier by the preparation of rating scales that focus attention on the salient aspects of any alternative learning experience. If the school district has measurement-trained staff members, their assistance should be sought in developing these rating scales. If no such help is available within the district, the assistance of staff at regional centers or other outside agencies would be highly desirable.

REFERENCES

1. Knapp, J. A guide for assessing prior experience through portfolios, CAEL Working Paper No. 6: Princeton, N.J.: Educational Testing Service, April 1975.

2. Forrest, A. A student handbook on preparing a portfolio for the assessment of prior experiential learning, CAEL Working Paper No. 7: Princeton, N.J.: Educational Testing Service, April 1975.

Address inquiries to CAEL, Educational Testing Service, Princeton, N.J. 08540.

CHAPTER 9

PERFORMANCE ASSESSMENT

> The scene is a high school music room. A student is playing the second of three piano selections, each of which represents a different musical period and style of composition. She is being evaluated by three judges: a much-respected organist from one of the area churches, a former student of the high school who is an advanced student of the piano at a music conservatory, and the student's faculty adviser, a music teacher at the high school. Although the student has made some mistakes and has had occasional difficulty maintaining an appropriate tempo, her performance meets the standards set by the school, and she will receive credit.

Definition

Performance assessment refers to a process in which the student is asked to carry out clearly defined tasks before at least two and preferably three or more judges under a given set of conditions. In performance assessment, criteria need to be specified that can be used to set acceptable performance levels. To carry out performance assessment, the school will have to have available appropriate materials or equipment on which the student can demonstrate the competencies that are of interest to the school. It is assumed that the student will be eligible for performance assessment only if a faculty adviser and the School Evaluation Committee have agreed that successful demonstration of a carefully specified set of competencies will constitute complete or partial fulfillment of credit requirements. It is best to reserve the performance-assessment approach for situations in which the student can demonstrate his or her ability to handle critical elements of important tasks.

A school will have to decide whether students will be asked to report on the training procedures that led to their acquisition of the skills under consideration. Such supplementary reports, which could be oral, tape-recorded, or written, will be appropriate if the performance assessment alone is not sufficient for making decisions regarding credit. In some contexts, such as equipment operating or athletic skills, the successful demonstration of an acquired set of competencies may be judged sufficient to satisfy credit requirements without any supporting information about how the competencies were acquired. A school would thus be saying: "We don't care how the student happened to learn these skills; it is enough to know that he or she has them."

Description of the System in Operation

As with the previously described methods of testimony and committee review, policies related to performance assessment covering such things as provision of information, eligibility, and application procedures will be developed by the school.

When a student petitions for performance assessment, he or she will have already ascertained through consultation with the faculty adviser whether or not appropriate equipment, materials, and supplies for demonstrating his or her competencies are available on site at the school. If no such equipment is available, special arrangements will be made for the student to be tested under controlled conditions outside the school setting. It may prove necessary to identify certain periods during the school year when students can be scheduled for performance-assessment activities. In each performance-assessment situation, the judge or judges who will evaluate the student's performance should be provided with a carefully detailed set of instructions regarding the conditions under which assessment is to be carried out, the attributes of the student's performance that are to be evaluated, and the acceptable level of performance.

The student who reports for performance assessment will have received in advance of the testing session a description of the kind of performance task that he or she can expect. The school will have to give careful attention to the level of specificity it provides to the student. A student of auto mechanics, for example, might be told that he or she would be asked to identify the cause of a malfunctioning carburetor and to repair that malfunction in a specified period of time. It will usually not be appropriate, however, to indicate to the student in advance of the assessment session precisely what type of difficulty he or she can expect. What needs to be recognized here is that the judge will be using the results of a particular performance to generalize to a larger body of potential problems in which the student is expected to have competency. Only in the situation where the performance assessment will cover all of the skills that are of interest can the student expect to receive a complete listing of each of the tasks in detail that he or she will be asked to carry out.

In addition to preparing standard sets of instructions for both the observer or examiner and the student, the school staff should develop a rating and scoring form for use with each student. Each performance assessment situation will require a specially designed set of forms and instructions. Appendix H provides examples of instructions to the observer or examiner, instructions for the student, and a rating form for a "troubleshooting" task that involves a defective generator.

Ideally, training in the use of rating or scoring forms would be carried out under supervised conditions in a pilot study. This would provide an opportunity for judges to reach agreement on criteria that would be applied in the actual performance-assessment settings and also to determine the degree of consistency among their ratings. The method of selection and the qualifications for judges will have to be established by the School Evaluation Committee. When choosing such judges, it will be useful to obtain recommendations and evaluations from individuals in similar lines of work. Community judges will need some exposure to typical high school-level performance so that they can avoid the tendency to apply the standards they might use with experienced workers in the field.

Possible Applications of the Performance-Testing Approach

Auto mechanics has already been alluded to as one possible area where students might be asked to demonstrate proficiency. Any school entering into assessment in the area of auto mechanics could profit from a review of the work associated with the Plymouth Trouble-Shooting Contest. An article by Bill Hartford in the January 1974 issue of Popular Mechanics outlines some of the procedures followed in the national finals of this contest. Pairs of students attempt to diagnose and repair identical automobiles each of which has seven different malfunctions built into its engine. The selection of contest winners is based on accuracy and the time required to repair a car as well as the results of a written examination on trouble-shooting skills.

Clearly, the individual school will seldom be able to provide a complete array of cars with carefully built-in malfunctions for their students to diagnose and repair. It will be possible, however, to build on the careful attention given in the Plymouth Automobile Repair Contest to identifying malfunctions that automobile engineers have judged reasonable bases for differentiating among carefully trained students. There are many other possible opportunities to use performance assessment in the industrial arts.

Another area in which performance testing has been applied and one that has been explored in detail is that of typing and related secretarial skills involving the use of office equipment. It is often possible to work with local business concerns to identify tests that are used in their selection and placement programs. As more and more students acquire skills with calculators and data processing-related equipment, schools can expect students to be petitioning for credit for these skills.

Although performance assessment is often thought of as necessarily related to the use of machinery, it is clear that demonstrating the ability to communicate in a foreign language can be a realistic and appropriate performance test. The Peace Corps, for example, has used language-proficiency interviews with its volunteers for many years and, in that connection, makes use of the following performance categories:

> Category 1: Able to satisfy routine travel needs and minimum courtesy requirements. Can ask and answer questions on topics very familiar to him; within the scope of his very limited language experience can understand simple questions and statements, allowing for slowed speech, repetition or paraphrase; speaking vocabulary inadequate to express anything but the most elementary needs; errors in pronunciation and grammar are frequent, but can be understood by a native speaker used to dealing with foreigners attempting to speak his language. While elementary needs vary considerably from individual to individual, any person at the Category 1 level should be able to order a simple meal, ask for shelter or lodging, ask and give simple directions, make purchases, and tell time.
>
> Category 2: Able to satisfy routine social demands and limited work requirements. Can handle with confidence but not with facility most social situations including introductions and casual conversations about current events, as well as work, family, and autobiographical information; can handle limited work requirements, needing help in handling any complications

or difficulties; can get the gist of most conversations on non-technical subjects (i.e. topics which require no specialized knowledge) and has a speaking vocabulary sufficient to express himself simply with some circumlocations; accent, though often quite faulty, is intelligible; can usually handle elementary constructions quite accurately but does not have thorough or confident control of the grammar.

Category 3: Able to speak the language with sufficient structural accuracy and vocabulary to participate effectively in most formal and informal conversations on practical, social, and professional topics. Can discuss particular interests and special fields of competence with reasonable ease; comprehension is quite complete for a normal rate of speech; vocabulary is broad enough that he rarely has to grope for a word; accent may be obviously foreign; control of grammar good; errors never interfere with understanding and rarely disturb the native speaker.

Category 4: Able to use the language fluently and accurately on all levels normally pertinent to professional needs. Can understand and participate in any conversation within the range of his experience with a high degree of fluency and precision of vocabulary; would rarely be taken for a native speaker, but can respond appropriately even in unfamiliar situations; errors of pronunciation and grammar quite rare; can handle informal interpreting from and into the language.

Category 5: Speaking proficiency equivalent to that of an educated native speaker. Has complete fluency in the language such that his speech on all levels is fully accepted by educated native speakers in all of its features, including breadth of vocabulary and idiom, colloquialisms, and pertinent cultural references.

Whenever performance assessment is used, it will be necessary to differentiate between knowing about a performance and being able to do it. Both the automobile troubleshooting and the foreign language proficiency examples described above illustrate this approach.

Strengths and Limitations

Because of the recognized limitations of paper-and-pencil testing for skills assessment, performance assessment is often overrated by individuals who lack experience with this approach. It is true that performance testing can avoid some of the obvious biases of essay testing in which the skillful user of language can often appear to know much more than he or she actually does about a given subject. Performance assessment, however, raises some problems of its own. To successfully plan and develop a performance-assessment program, it is first necessary to identify those skills that are of critical importance in any training area. If the training program is carried out within a school, one can turn to those members of the school staff who have expertise in the tasks under consideration. It will often be the case, however, that students will wish to demonstrate their proficiency

at tasks that require competencies the school staff do not possess. Training materials may prove helpful as a source of information about which skills are the critical ones for certain tasks. However, it will usually be necessary to consult with individuals in industry about which elements of a task constitute an appropriate basis for performance assessment. (See Sharon, A Task-Based Model for Assessing Work Experience in the reference list at the end of this chapter for suggestions on how to accomplish this task.) As with any testing approach, the student will only be able to demonstrate his or her competency on a sample of the tasks that might be posed. It is critical, therefore, to make that sample as representative as possible of the performance area that is of interest.

Establishing the conditions under which performance assessment will be carried out will require a good deal of staff attention. Clearly some of the malfunctions in the Plymouth carburetor mentioned earlier will be much more difficult to diagnose than others. Identifying only one particular type of malfunction as the sole basis for testing would not be desirable since students could then become proficient at that task alone, and the performance-assessment procedure would no longer be adequate for generalizing to other competencies that are of interest.

School staff should guard against the tendency to vary the conditions for assessing competency from one student to another. For example, a school may be uncertain about whether or not to give its students a manual of instructions for carrying out a certain fundamental operation involving a machine. If the school decided to give a manual to its least proficient students, comparability of assessment from student to student would be destroyed.

Schools would be well advised to consider the prerequisites that they will ask of a student before approving any performance assessment credit-granting arrangement. It will always be important to ensure the safety of the student, a particularly important issue when dealing with electrical work and heavy machinery. In addition to a concern for safety, schools should require students to demonstrate readiness for performance testing. If, for example, the school recognizes that work with certain types of electronic equipment requires some knowledge of mathematics, students should be encouraged to acquire the necessary mathematical skills before learning to use this equipment. The school could use preliminary screening tests to determine whether students are indeed ready for performance training. Many tasks that seem on superficial inspection to require only manual performance actually can only be accomplished successfully after the student has read a good deal of technical material that requires a fairly high reading level. If there is no direct instruction that could take the place of reading the technical material, the student should be screened and prevented from entering the program until he or she has demonstrated the ability to handle the required reading.

Technical Issues

The preceding discussion has identified the task of preparing explicit guidelines for both the student and the examiner as background for any performance-assessment procedure. A student workbook outlining the procedures

that will be followed during the assessment will be an essential guide to keeping a program running smoothly. The school staff will, however, have to review the information they are planning to give to students and undertake the difficult task of trying out different assessment approaches. It may be possible to use certain performance-assessment procedures on an experimental basis in a regular program within a school. Various performance-assessment procedures could be developed on an experimental basis and administered to those students who have completed a formally supervised program.

When a student asks to carry out a learning program to receive credit for acquiring job-related performance skills that are not part of the regular school program, every effort should be made to draw on the resources of local industry, colleges, or regional vocational-technical schools where curriculum experts or expert practitioners can identify appropriate criteria and conditions for making judgments.

As with any assessment procedure, teachers, administrators, and cooperating community members need to guard against the possibility of setting unrealistically high standards of performance for securing credit. A typing teacher, for example, may wish to require students to demonstrate a typing proficiency of 80 words per minute, hoping to insure that anyone who meets this criterion would be able to reach a practical criterion of 60 words per minute on the job. It may actually be true, however, that given the kinds of interruptions and competing activities that the working secretary faces, a typing speed of 40 words per minute under usual office conditions may be quite adequate. Wherever possible, schools should check with other schools and local industries for performance data on individuals who are actually carrying out in the real world the kinds of tasks for which students are asking to be assessed. Standards can then be established that conform to the entry-level requirements of individuals who work with various types of equipment. The performance-assessment system can then be designed and developed to be realistic and fair to the students who are seeking credit through experiential learning.

REFERENCES

Boyd, J.L., Jr., & Shimberg, B. Developing performance tests for classroom evaluation. Princeton, New Jersey: ERIC Clearinghouse on Tests, Measurement, and Evaluation, ERIC TM Report No. 4, June 1971.

Boyd, J.L., Jr., & Shimberg, B. Handbook of performance testing, a practical guide for test makers. Princeton, New Jersey: Educational Testing Service, 1970.

Fitzpatrick, R.A. & Morrison, E.J. Performance and product evaluation in Thorndike, R.L. (ed.) Educational Measurement (2nd ed.). Washington, D.C.: American Council on Education, 1971, 237-270.

Hartford, B. Plymouth's Annual Trouble-Shooting Contest: This race really is "won in the pits"! Popular Mechanics, January 1974.

Knapp, J. & Sharon, A.T. A compendium of assessment techniques. Princeton, New Jersey: CAEL, Educational Testing Service, 1975.

Panitz, A. & Olivio, T.C. Handbook for developing and administering occupational competency testing. New Brunswick, New Jersey: Rutgers University, 1971.

Sharon, A.T. Planning the development of measurement and evaluation services for use in occupational programs at postsecondary institutions. (PR-74-16) Princeton, New Jersey: Educational Testing Service, 1974.

Sharon, A.T. A task-based model for assessing work experience, CAEL Working Paper No. 8. Princeton, New Jersey: Educational Testing Service, 1975.

Organizations that may be of assistance include:

Cooperative Assessment of Experiential Learning (CAEL)
See references in Chapters 6 and 8.

Clearinghouse for Applied Performance Testing (CAPT)
The Clearinghouse for Applied Performance Testing (CAPT) was created in 1974 by a consortium of state departments of education. CAPT collects and develops information and instructional materials related to the construction, use, and evaluation of applied performance testing. Applied performance testing is defined by CAPT as:

> "the measurement of performance of some task significant to a student's life outside the school and/or to adult life. Such a task is valued as output for public schools. The testing device must allow for measurement of the task in an actual or simulated performance setting"

Schools can request free placement on the CAPT mailing list and receive further information by writing to: Director, Clearinghouse for Applied Performance Testing (CAPT), Northwest Regional Educational Laboratory, 710 W. Second Avenue, Portland, Oregon 97204.

National Occupational Competency Testing Institute (NOCTI) Examinations
A series of written and performance examinations in 24 major occupations administered nationwide on a regular basis. The major purpose of the examinations is to grant academic credit to individuals with work experience who desire to become vocational teachers. For information write to: Program Director, NOCTI, Educational Testing Service, Princeton, New Jersey 08540.

CHAPTER 10

PRODUCT ASSESSMENT

> The scene is a meeting room of the high school. A topographical map of the city rests on a table. The six high school seniors who worked together to prepare the map are describing their role in the project to three judges: a member of the city zoning commission, a professional mapmaker, and the students' faculty adviser, a social studies teacher. "I collected the reference material for all of us." ..."I drew the topographical lines." ..."I cut the templates and etched the lines." ..."I lettered the names." Since all the students have satisfactorily carried out their individual tasks in the planned joint project, all will earn credit.

Definition

Product assessment as used in the context of evaluating the competencies of high school students refers to the evaluation of a variety of student products such as maps, drawings, paintings, musical compositions, science projects, diaries or reports of travels, photographic essays, and handcrafted items such as cabinets or bookcases. Both individual and joint projects can be planned and assessed. When a joint project is approved, however, the role of each student will need to be defined. Joint projects can provide opportunities to include students with special needs. The contribution of such students to appropriate integrated group activities can be defined in accordance with their strengths and abilities.

Product assessment can be used whenever the tangible results of a student's efforts constitute appropriate evidence for judging the quality of his or her learning experience. School staff members will want to seek help from community members when making decisions about what kinds of products to assess and which aspects of each product to evaluate.

Product assessment differs from performance assessment in that the latter approach can be used whenever it is possible to observe the student developing or creating the finished product. Both process and product are evaluated in performance assessment. The distinction between the two approaches becomes less clear in situations such as performing a musical composition, when the student's performance might be observed initially but replayed later on tape for the benefit of other judges from the school and community.

The product-assessment approach can be used as an indirect way of evaluating student learning and acquisition of skills that have taken place outside the school setting. A student will be eligible for credit through the product-assessment approach only if the School Evaluation Committee has agreed that an acceptable level of quality of the particular product or products constitutes complete or partial evidence toward the granting of credit.

Careful planning and monitoring of the students' product-development program should make it possible for almost all students who choose this alternative learning and assessment approach to earn credit for their work. It should be recognized, though, that a school that uses this approach will need to consider strategies for verifying and certifying that the student is indeed the author or developer of the materials to be judged. One method that can be used is oral questioning regarding the materials and procedures. If extensive questioning is used, however, schools should recognize that the product-assessment method will become a particular type of committee review, as described in Chapter 8.

Description of the System in Operation

As with each of the other assessment methods recommended in this Handbook, policies covering elements such as provision of information, student eligibility, and application procedures will be developed by the School Evaluation Committee.

Preliminary consultation with a faculty adviser should be an integral part of the product-assessment system in a school. The student will be required to demonstrate to the satisfaction of the faculty adviser that he or she will be engaging in appropriate learning activities in the course of developing the required products. In many circumstances, it is likely that product assessment will be only a part of the competency-assessment procedures that lead to the award of course credit. Rarely will a single product or set of products be adequate to grant full credit in the absence of appropriate supporting information. In addition to conducting an initial interview with the student to determine the learning activities he or she is planning, the faculty adviser should schedule additional meetings, as necessary, to monitor the student's progress. These periodic meetings should indicate whether the final products will be able to stand alone for examination by school and community judges or whether the student will be able to produce an appropriate supporting document to explain the nature of the product and thus provide the judges with whatever supplementary information might be needed for them to make their decisions. A critical question the faculty adviser and the School Evaluation Committee will have to answer is: Without having seen the development process, can the judges consider the products themselves as sufficient evidence of accomplishment?

The qualifications of the judges will have to be established by the School Evaluation Committee. These qualifications can be expected to vary according to the nature of the product submitted. Whenever possible, the judges should be individuals who have had prior experience judging the work of other students who have used the same materials and approach. The school should facilitate the work of the judges by developing rating forms that specify which aspects of a product should be judged. It is easiest to develop rating forms when a number of students are required to submit a standard set of products, such as a portfolio of artistic work, specific woodworking objects, or other products requiring the application of the student's skills toward the development of a prespecified set of final products.

The judges should use a separate rating scale for each of the critical dimensions of the product or products under consideration. In most circumstances, a scale with five to seven points representing different levels of quality should be adequate. (Although it is possible in some circumstances to use scales with a much larger number of points, it is very difficult to make differentiations and discriminations when scales of 15 or 20 points are used.) Judges should work independently in arriving at the scale value for each of the attributes of the product under consideration. The final decision whether a student's work has reached a minimal acceptable level should be reached by combining the judges' ratings on each attribute and arriving at a total score, which can be compared against previously established standards using the methods described in Chapter 5. In addition to rating each significant attribute of each product, some consideration should be given to rating the total set of materials submitted by any student.

Possible Applications of the Product-Assessment Approach

One area where product assessment has a fairly long history of successful use is in the evaluation of artistic products. Schools interested in applying this approach to the artistic works of students should obtain copies of the document entitled Evaluating the Advanced Placement Studio Art Portfolio (1). This document describes the preliminary planning and implementation phases of a national program that evaluates a portfolio of students' artistic works. The procedures employed in the Advanced Placement Program are designed to determine whether or not a student has successfully completed college-level work in the artistic area. The same procedures can be applied to evaluating whether or not a student has adequately demonstrated high school-level work.

When applying the product-assessment approach to artistic, scientific, or other student-developed materials, the faculty will have to decide how much weight is to be given to the learning that has occurred and how much to the quality of the final work. In some situations, it will be appropriate to require that the work produced be of high quality and to withold credit unless this level of quality is attained. It seems reasonable to impose this requirement, however, only in situations in which the faculty adviser has determined that the student is capable of producing high quality work. When the learning experience itself is of primary interest and the school is prepared to offer credit even when the final work is not of high quality, it would be inappropriate to use the product-assessment method of student evaluation. For example, a student might attempt to create a photographic essay and learn a great deal about the limitations of types of equipment and the proper methods of using them. The school could award credit for this learning through the committee-review assessment procedures, even though the student was not able to produce individual photographs of professional quality. Conversely, a student whose learning contract emphasized the technical aspects of photography could be given credit for taking and developing good photographs even if he or she were not able to tell a story with them.

Strengths and Limitations

The product-assessment approach will be a valuable component of a school's competency-evaluation activities when individuals can be located who have the necessary skills and training to permit them to judge the kinds of student works that the students themselves are interested in producing. If faculty members have opportunities to evaluate the students' works and find it possible to agree on which aspects of the work should be evaluated and what level of quality would be considered acceptable, the primary tasks facing the school are to develop and monitor an appropriate management system for the assessment method. Substantial difficulty can be expected, however, if school staff lack experience with the kinds of products that students would be most interested in submitting for consideration for the award of credit. The use of appropriate members of the larger community can be of assistance here as long as adequate provision is made for training in the field of judging student work. It will sometimes be possible to locate licensed professionals in a field who will be willing to serve as judges, but it will often be necessary to rely on the evaluations of other community members in the same field. Although community members have considerable potential for contributions to the evaluation process, their efforts will be of limited value unless the school can make arrangements for professional assistance in the development of appropriate rating scales for students' work.

What is needed is a product-assessment system that has clearly defined guidelines. The student can use these guidelines in preparing materials that will be examined under controlled conditions and with carefully focused attention to the critical aspects of his or her work. A set of procedures for monitoring the adequacy of judgments will be essential, and the School Evaluation Committee will need to provide for a periodic review of the kinds of standards that are being employed by school and nonschool judges.

Technical Issues

The discussion in Chapter 9 of technical issues related to performance assessment also applies to the area of product assessment. In both instances, a student workbook should be developed that contains a description of the assessment procedures that will be employed. The student should know the number and type of works that need to be submitted, and the judges need to know precisely what their responsiblities will be in connection with each product-assessment session. In many circumstances, it will be desirable to study the works produced by students in the school's regular programs and to identify models of unacceptable, acceptable, and particularly high quality works that judges can use as standards against which to compare any new product. In some instances, the actual products could be maintained for reference use by the judges. In other circumstances, it might be more convenient to store and make available photographs of the materials.

In schools where product assessment is being introduced as an approach to evaluating student competencies, it is recommended that the school carry out pilot or experimental work in one or two subject matter areas. These experimental projects should develop and refine procedures for describing the works to be created and for setting up and conducting a product-assessment session. When planning the experimental project, it will be useful to refer to reports outlining significant issues and procedures (2,3). Professional assistance should be sought in the development of rating scales and in evaluating the consistency of ratings from judge to judge for different products. Since studies of the rating of both essays and student work have demonstrated the problems associated with obtaining such ratings, it is essential to provide adequate training for judges, preferably in a workshop setting involving samples of actual students' work. In this connection, the works produced by students enrolled in appropriate courses would provide good training material for the judges to use in deciding which aspects of the work should be judged and the level of quality that should be considered acceptable for the purpose of awarding full or partial credit to a student.

REFERENCES

1. Dorn, C. M. Evaluating the Advanced Placement Studio Art Portfolio. Princeton, N.J.: College Entrance Examination Board and Educational Testing Service, 1974.

2. Fitzpatrick, R. & Morrison, E. J. Performance and product evaluation in R. L. Thorndike (Ed.) Educational Measurement (2nd Ed.). Washington, D.C.: American Council on Education, 1971, pp. 261-262.

3. Knapp, J. & Sharon, A. A compendium of assessment techniques. Princeton, New Jersey: CAEL Educational Testing Service, 1975.

CHAPTER 11

OBJECTIVE AND ESSAY EXAMINATIONS

> The scene is a high school classroom. Ten students are sitting for examinations that could permit them to earn credit. This particular group includes students who are taking biology, English, and geometry. Within a week, the students will learn whether they have earned credit. Those failing to earn credit will need to show evidence of further study if they want to take an examination for possible credit next semester.

Definition

Objective and essay examinations constitute an approach to assessment in which a formal set of examinations is used to assess student competencies. The examinations should be reviewed and selected by a school or district-wide committee or designed and developed with the help of such a committee. Moreover, student responses to the examinations should be in either a written or tape-recorded form so that scoring can be done after the test administration under controlled conditions.

The discussion in this chapter is addressed primarily to paper-and-pencil tests. Such tests are now widely used nationally to permit high school students and adults to earn college credit via the College Entrance Examination Board's Achievement Tests; the Advanced Placement tests; and the College-Level Examinations. State and institutional examination programs also offer opportunities for earning college credit. The programs that focus on college credit have some features that merit consideration when considering the use of examinations to award credit at the high school level. It should be noted, for example, that the tests used in such programs are linked closely to the content of particular types of college courses. At the high school level, it will be appropriate to select or develop tests covering the content of particular secondary school courses. Some use might also be made of tests covering broad areas such as humanities or social sciences.

Description of the System in Operation

As with each of the other assessment methods reviewed in the Handbook, policies must be developed covering provision of information, student eligibility, application procedures, and the like. The chapter of this Handbook on management standards outlines procedures that could be followed to define policies and set procedures. Some topics are reviewed on the following pages that will need special attention if objective and essay examinations are used to assess student learning.

Credit and/or Exemption: The Handbook is based on the assumption that examinations will be used primarily to make decisions regarding awarding of credit. If students have acquired the knowledge and skills covered by the examination, they can expect credit for the course. Some schools, however, may be interested in establishing an examination system that permits students earning scores below the standard for credit to nevertheless be exempted from taking what would otherwise be a required course. The examination-system approach can be used by schools wanting to offer exemption only, a choice of exemption or credit depending on students' performance, or credit only.

Awareness: A list of courses for which credit-by-examination may be earned will need to be published and made available to students.

Eligibility: Schools may want to restrict the credit-by-examination option to students who have completed one or more years of a standard program. It might also be desirable to delay until the junior or senior year the opportunity for students to challenge health or social studies/citizenship courses. This approach might be taken if it seemed appropriate to limit the credit-by-examination option in these areas to students who had acquired somewhat greater maturity.

Application Procedures: As in the case of the other assessment methods, a clear definition of the application procedures should be made readily available to students. The procedures should specify whatever approvals from parents, faculty advisers, and others were necessary and should indicate any cut-off dates that would exist for students to apply for examinations. Clearly, students who will have to take the courses in the regular program if they fail to achieve credit through the examination route will need to take their examinations early enough so that it will still be possible to enroll for the required courses.

Security Restrictions: One critical issue will be the need to maintain the security of the tests that are to be used as a basis for awarding credit. In addition to arranging for the safe storage of examinations and for restricted access to them, some constraints will be necessary on the number of times an examination may be taken and on the length of time that must pass before retesting. More than one form of the test will be needed and it may be useful to limit the number of times a student can take a test to the number of different forms that are available.

Special Restrictions: Schools may wish to limit the student to two opportunities to challenge a course through the examination route and to specify that the attempts must be made in different semesters. Schools may want to require the student to demonstrate some additional evidence of a learning program before being allowed to retake an examination in an attempt to win credit.

Policies should include the provision that no credit-by-examination could be earned for a course lower in a sequence of courses than one that has already been completed by the student.

The policies might bar the student from petitioning for credit-by-examination for a course in which the student is currently enrolled after the first three or four weeks.

The total number of credits that can be earned by credit-by-examination will likely be limited to 25 or 30 percent or some such proportion of all credits needed for graduation.

Test Selection or Development

In addition to having a school evaluation committee that considers overall policy and strategy related to the award of credit to students, it will be essential to have a committee responsible for the selection or development of tests. Clearly, the School Evaluation Committee and the test-selection or development committee will need to have overlapping memberships. In those instances in which the primary assessment model to be employed by the school is credit-by-examination, the committees can be identical. It should be recognized, however, that the test-selection or development committee must evaluate the appropriateness of test content. Thus, its membership should include subject matter specialists in the fields for which exams are to be developed. A number of the major issues that the committee will need to address are treated in the following sections of this chapter.

Should our district use existing examinations or develop new ones? It is important to recognize that there are issues of social values associated with the choice of either an external or a locally developed test. An external examination will often have greater credibility with the community, but it may be viewed less favorably by students and teachers than a locally developed examination. It will be essential to involve all groups in the decisions regarding a test-development or selection strategy.

When considering the use of existing externally developed standardized examinations, school districts are likely to react positively to the following features of such examinations:

- Such examinations will typically cost less than locally developed examinations.
- High technical quality may be obtainable. Careful reviews of publishers' manuals and test copies will be helpful in evaluating technical quality.
- The job of setting up an assessment program that uses externally developed examinations is likely to make fewer demands on the time of the school staff.
- External examinations will usually be available in multiple forms. This reduces considerably the problem of test security.

These positive features of externally developed examinations need to be weighed against such negative points as the following:

- No available examination may match adequately the objectives of courses for which the district wants to offer the credit-by-examination option.

- Existing examinations may have a cultural flavor that is inappropriate for a significant number of potential test takers.

- The standardized administration approach may be inappropriate for some students such as those with visual handicaps or learning disabilities. If such students have indeed mastered the course objectives, some assessment procedure that permits them to demonstrate this fact will be essential.

- Lack of widespread acceptance of the appropriateness of external examinations would undermine any credit-by-examination system regardless of the technical adequacy of the actual instruments.

How can our district obtain information about existing examinations? A comprehensive approach to the problem of searching for background information on tests is suggested below. If several of the recommended major sources are used, it should be possible to collect a great deal of information about existing tests. The discussion focuses on three categories of procedures for a systematic search for tests and assessment devices:

I. Preparatory Activities
II. Literature Search
III. Contacts in the Field

I. Preparatory Activities

For a search to be productive, adequate time must be spent first in deciding what is needed. Some important steps to consider when contemplating a search for information about tests or other measurement instruments are:

A. What learning outcomes or competencies need to be measured? What type of test is needed and how will it be used? What decisions regarding credit and/or exemption will be made and what part can test-based information contribute to these decisions?

B. What are the constraints that will influence the search and subsequent test development?

1. Time and Schedule--When is the test needed? How much time can be devoted to the search and to the task of development?

2. Available Personnel--Will a single staff member do the search or are several people available? If more than one person will help, consideration should be given to ways of dividing responsibility.

3. Fiscal Considerations--Can the district afford to purchase materials about tests or to design a testing approach that requires expensive equipment?

4. Nature of Population--Who will be tested? Consider such factors as age, previous familiarity with tests, special needs that will have to be taken into account, language development (particularly if bilingual students are to be tested), and motivation to cooperate.

5. Available Facilities--What search sources are convenient such as a nearby college library? What sources can be reached only with considerable effort?

II. Literature Search

The search will be greatly assisted if school staff can locate a recent book or article that summarizes much of the recent work in the area of interest. The searcher should identify for further study the references that relate most closely to the district's interest and determine what journals would be likely to publish additional articles relevant to the project.

The task of reviewing a field is best handled by referring to one or more of the summaries that are available. The searcher can learn what tests exist and read critical reviews of these tests.

A. Mental Measurements Yearbook Series

1. Mental Measurements Yearbooks
2. Tests in Print
3. Reading Tests and Reviews
4. Personality Tests and Reviews

This Mental Measurements Yearbook series (Gryphon Press, Highland Park, New Jersey) includes descriptions of tests, critical reviews, publishers' directories, and bibliographical references.

B. CSE/Secondary School Test Evaluations

These three volumes contain ratings of available secondary-level tests on a number of different factors. The Secondary School volumes cover grades 7-8, 9-10, and 11-12. For additional information write to:

Dissemination Services
Center for the Study of Evaluation
Graduate School of Education
University of California
405 Hilgard Avenue
Los Angeles, California 90024

C. NCME Measurement News

This newsletter of the National Council on Measurement in Education contains general articles on testing issues as well as announcements of new tests and lists of test reviews.

D. Test Collection Bulletin--ETS, Princeton, New Jersey

This is a quarterly digest of information on tests and services that generally have become available after the publication of the most recent Mental Measurements Yearbook. Commercially available and research measures are described. The Bulletin does not evaluate the tests that are listed but it does provide references to test reviews.

E. Promotional Materials from Test Publishers

Check publishers' catalogs and announcements for references to tests, services, and technical data on specific measures made available after the publication of the Mental Measurements Yearbook.

F. An Annotated Bibliography of References to Tests and Assessment Devices

This publication may be obtained without charge from the ETS Test Collection, Educational Testing Service, Princeton, New Jersey 08540.

G. Educational Resources Information Center (ERIC)

ERIC is actually a series of clearinghouses, each specializing in a specific area of education. Reports and articles are collected and indexed in ERIC publications: 1. Research in Education, 2. Current Index to Journals in Education, and 3. Clearinghouse Publications (frequently a source of state-of-the-art papers). Items cataloged in the ERIC system can be ordered in microfiche or hard copy. The Merrimack Education Center maintains an extensive collection of ERIC documents. The ERIC Clearinghouse on Tests, Measurements & Evaluation is located at Educational Testing Service in Princeton, New Jersey.

H. Professional Journals

Professional journals such as The Reading Teacher and the Journal of Educational Measurement provide not only direct reports of developmental and research activities but also related theoretical papers. The advertisements in the journals are an excellent source of information on new materials and services available from commercial publishers and research organizations.

I. Traditional Reference Tools

The following sources of abstracts of educational and psychological

literature can usually be found on the reference shelves of college libraries:

1. Psychological Abstracts
2. Education Index
3. Research Studies in Education
4. Dissertation Abstracts

III. Contacts in the Field

A. Information Directories

Particularly useful is the Encyclopedia of Information Systems and Services, edited by Anthony Kruzas (New York: Edward Brothers, Inc., 1971).

B. Educational Testing Service Test Collection

This extensive library of tests and other measurement devices also includes publishers' catalogs and descriptive materials, information on scoring services and systems, test reviews, and reference materials on measurement and evaluation. It provides on-site, telephone and mail reference services. Districts can write for lists of publications that summarize available instruments in specified areas (Test Collection, Educational Testing Service, Princeton, N.J. 08540).

C. Massachusetts Department of Education

The Division of Research and Evaluation can provide information regarding the state assessment program. Other divisions can supply information regarding particular subject matter fields or technical approaches.

D. Research and Development Centers and Child Study Laboratories, etc.

E. Professional Organizations and Special Interest Groups such as

American Printing House for the Blind
1839 Frankfort Ave., P. O. Box 6085
Louisville, Kentucky 40206

International Reading Association
800 Barksdale Road
Newark, Delaware 19711

American Association for Health, Physical Education and Recreation
1201 16th Street N. W.
Washington, D. C. 20036

How Should New Objective and Essay Assessment Examinations Be Developed?

The examinations used for awarding credit and the ways in which they are developed are likely to receive a great deal of critical attention from educators, parents, and others. It is important, therefore, to maintain high measurement standards in their design and development. This goal will most likely be achieved if schools use staff members who have training in measurement and statistics. A school district work group can include some relative newcomers to the field of testing, but it must have a solid core of technically trained staff.

Any testing project is likely to succeed or fail on the basis of the quality of the staff involved, yet even a good team is not sufficient. The odds that good staff will do a good job will be heavily influenced by the extent to which adequate planning takes place. On the following pages are

areas of test development that will require careful thought. Each will present its own special problems. The points to be considered are grouped into the following seven areas:

Initial Planning and Allocation of Responsibility
Development of Examination Specifications
Test Question Development
Pretesting
Use of Statistical Item Analysis
Final Test Assembly
Administration and Scoring

Each of these is considered in turn.

Initial Planning and Allocation of Responsibility: 1) Planning should begin with the identification of all components of the examination development project--This step involves the participation of all project staff, possibly supplemented by external consultants with skills that round out talents of the project team. An extremely useful source of information for this and other issues related to test development is the chapter "Planning the Objective Test" by Sherman Tinkelman in Educational Measurement edited by Robert L. Thorndike (8).

2) Development of a schedule for the completion of the steps--This task is most readily accomplished by working backward from the time when examinations will be needed. The length of time needed to accomplish each step is determined using whatever sources are available. The identification of critical sequences can often be facilitated through the use of flow charts or other diagnostic or tabular methods of presenting data.

3) Establishing clear lines of responsibility--A coordinator should be designated who can assign responsibility for aspects of the work to members of the work team on the basis of their experience and competencies. A written statement of responsibilities will be useful for large working groups.

Development of Examination Specifications: 1) The specifications for the assessment examinations should be developed with the help of a community advisory group representing many sectors of the community. This group should be involved in early project planning. The longer one waits to involve an external group in a project, the more likely the members will resent the possible implication that they are being called in to rubber stamp the plans of school staff. Because it is difficult to make changes late in the process of examination development without bypassing desirable review and quality-control steps, school staff members are likely to resist suggestions for change. An advisory group that is involved early in the development process can help with both easy decisions on which there is consensus and on difficult ones that can be handled only through compromise. An advisory group that has worked through this process will be more likely to defend than to criticize the resultant specification.

2) Covering all types of specifications--Textbook discussions of test specifications often center narrowly on the subject matter content of cognitive tests. It is important, however, to consider many different types of specifications.

a) Statistical Specifications--Appropriate statistical specifications or selection criteria for individual questions and for sets of questions need to be developed. Question difficulty, as judged by the percentage of students answering correctly, can serve as a check on the reasonableness of the content of the test. Teachers sometimes ask questions that even the very best students cannot answer. The correlation between individual questions and the total test will be useful as an index of item homogeneity and as a stepping stone to the evaluation of score reliabilities. Since reliability indices permit an estimate of the likelihood that a similar score would be earned on a parallel set of items, this information is essential to an adequate evaluation of any test.

b) Question Type Specifications--A number of practical constraints have led professional test developers to rely primarily on paper-and-pencil, machine-scorable question types. Each group developing an examination system needs to consider, though, the possibility that other approaches would be more appropriate to the subject under consideration. Consider, for example, the measurement of writing ability. Objectively scorable question types have been developed and validated against actual writing ability, yet it is clear that writing ability can only be measured directly through exercises requiring writing. Inclusion of actual writing exercises in a test does create such problems as the need for professional scoring of exercises, but the likely increase in the credibility of the test results may make it worthwhile to attempt to deal with the problems. (The issue of objective versus essay examinations is discussed in greater detail later in this chapter.)

c) Stimulus-material Specifications--In addition to reviewing the possibility that a variety of question formats might be feasible, attention should also be given to the use of other-than-written stimulus material for questions. Tapes, films, and slides might be employed, or test administrators can be trained to read certain materials, speak certain sounds, or make use of apparatus of various types. While budgetary factors must be taken into account, a credit-by-examination system must provide needed information if it is to be of value. Some types of needed information cannot be obtained by the least expensive testing formats.

d) Cultural Values Specifications--Cultural values are usually thought of as the province of some special area of testing, such as citizenship, if they are thought of at all. Yet tests do communicate values to students, and it is well to consider this fact at the planning stage. As noted above, school staff should be sure to represent various groups in the test-development process. It will also be helpful to develop guidelines regarding test content that will keep attention focused on an

appropriate balance of contributions from many different facets of subject matter fields. Throughout the test-development process, the work team should consider the cultural values that are implied by the selection of stimuli and the nature of the test questions.

e) Other Specifications--The foregoing specifications do not exhaust the list of question and total-test attributes to which the developers of examinations need to be sensitive. They do indicate the breadth of concern that is essential to successful program development. Test developers will need to identify the additional areas for which specifications will be needed.

Test Question Development: 1) Specifications first--Writing test questions is such a difficult, important, and time-consuming part of test development that there is a strong tendency to begin writing questions without adequate attention to detailed specifications. It is essential, though, to clearly specify what is to be measured before question writing commences. In some instances, this may require the elaboration of detailed educational objectives that are implied by, or subsumed under, existing course goals. Most school districts have staff members with training in the area of objectives writing, and there are many available texts on the subject (4,5,6,9).

2) Use of existing models--New test development is often initiated because of dissatisfaction with the questions on existing tests. It is inefficient, nevertheless, to ignore existing tests as a source of models, or at least ideas, for new questions or exercises. Much can be gained by collecting existing tests and taking a careful look at what is or is not desirable about the questions contained in them. A new test can use any format or approach that seems suitable while avoiding the undesirable features of existing tests.

3) Staff for question development--Every test development work group will need to decide who should write and review the needed test questions. Can a local school district obtain help for its program from the department of education, possibly through the Regional Education Centers? What part should outside experts from test publishers, research laboratories, or resource centers play in the process? These are questions that each school district must answer in the context of the available options. Whatever the direction taken, though, staff experienced in instrument development must play a major role in the question-development process.

4) Training question writers--One effective technique for training question writers is a "hands-on" workshop. A good workshop provides participants with training in the elements of successful question writing and incorporates a goodly amount of actual writing and reviewing experience. Generally, two or more days will be required so that two or three full cycles of question writing, review, and revision will be possible. The best actual writing of questions is usually done by question writers working independently. Question review and revision, on the other hand, should involve both individual and group work. The group sessions are opportunities to discuss different aspects of questions and to explore alternative approaches. The individual sessions permit the most efficient production of materials.

It is useful if participants in a question-writing workshop can be prepared in advance through the use of background materials that clearly define the questions-writing task and provide both terminology and the fundamentals of technique. (For assistance in preparing such material, see Ebel [2], Tinkelman [8], and Wesman [10] or similar discussions in testing textbooks.) If possible, trainees should even write some questions to the appropriate specifications and bring them to the first training sessions.

One component of a question-writing session should be the attempt to generate ideas for questions or exercises that can be further developed at a later time. This is a strategy that has been employed in the development of exercises for the National Assessment of Educational Progress (3).

5) Types of reviews needed--There are many reasons for reviewing test questions. Three of the most important are identified below:

- To be certain that the test questions are appropriate measures of the content to be covered, questions should be reviewed by individuals thoroughly familiar with the subject matter. Such reviewers can certify the appropriateness and accuracy of questions and with guidance from measurement-trained staff, can help evaluate the scorability of exercises that require judgment on the part of the scorer.

- No matter how many contributions subject matter specialists have made to question review, another review is always needed for consistency of style and clarity of expression. This review is best entrusted to a skilled specialist who has the same role for all examinations and questions. This procedure facilitates uniformity of format and style. If possible, this same editor/reviewer should also have responsibility for controlling the readability level of questions and of associated directions and explanatory materials. The examinations will be fair and useful only if students can understand the tasks posed to them.

- As a final suggestion in this area, developers should consider a possible review role for parents and other concerned citizens who are representatives of a crucial audience for test results. If parents are to contribute effectively, they should be brought into the review process early and should be given appropriate background information about the purposes and procedures of the credit-by-examination program. Members of the community advisory group can be helpful in this regard. Nontechnical reviews can serve a valuable public relations function and can bring helpful information on issues such as the importance ascribed to various types of course material and possible offensiveness or controversiality of exercises. Some input on these issues can also be obtained by giving students a role as reviewers. Students can be given an opportunity to comment on items as part of a post-pretesting session, when pretesting is included as a step in question and examination development.

When arranging for reviews by parents or teachers, it will be important to preserve the security of the examinations. Using numbered test copies that are collected after the review session will be helpful in this regard.

Pretesting: 1) Use some type of pretesting — Pretesting provides valuable information to the test development staff including additional protection against faulty questions. Even when time is limited, some form of question tryout is almost always possible. At a maximum, two strategies ought to be considered. The first involves the use of questions for the first year's examinations that have already been used in classroom tests or used in some other district's credit-by-examination program. The second strategy is simply that of carrying out analyses after the first use of the examinations and revising the examinations for subsequent years.

2) Development trials--One form of pretesting that requires very little in the way of time and money is a tryout conducted by the original question writer with a small number of students. In these circumstances, the questions can be administered on an individual basis and the students can be interviewed afterwards. This type of pretesting does not lead, typically, to the development of statistics. Rather, it permits an opportunity for the individual who is most familiar with the intention of the question to check the clarity of wording of questions or directions. The writer can observe what it is that the students seem to be doing when they answer or solve the problem or question. Do they appear to be carrying out the process originally intended by the question writer or are they using some other method of approaching the question? A question, for example, that is designed to require a student to use insight or to synthesize data from many sources would be judged suspect if students seemed to be answering solely from factual recall.

Developmental trials provide an excellent opportunity to discover vocabulary or phrasing of questions that is simply too difficult for the age level for which the questions are intended. To achieve the maximum benefits from developmental trials, they should be conducted with students drawn from the full competency range under consideration, including both very limited and quite highly skilled students. Another potential benefit from developmental trials is that one has an opportunity to obtain a first estimate of the amount of time students will need to respond thoughtfully to the test questions.

3) Small-group trial--Perhaps the next level of pretesting in terms of time and money required is the "quick and dirty" administration of questions and test materials to small groups of students. This can be done without preparing the same level of quality of test materials as is planned for the final administration. The small-group trials can include interviews with students in an effort to determine the effectiveness of the test questions in communicating the nature of the task. At a slightly higher level of involvement, the pretesting could incorporate sample questions from each of the various types of questions planned for inclusion in the final instruments. The questions chosen for pretesting should be representative of other questions, including some at the upper limits of complexity and difficulty.

4) Full-scale tryouts--Whenever possible, the test developer should try out questions with students as representative as possible of the final population of students. One way to accomplish this is to include a pretest or experimental section in the regular test.

Use of Statistical Item Analysis: 1) Question difficulty--As was noted earlier, information about the difficulty of questions is valuable for test development. Whenever questions can be scored right or wrong, as is the case with most multiple-choice questions in the cognitive domain, question difficulty can be determined. Similarly, one can determine the difficulty of sets of questions such as all the questions related to a single objective or all the questions relating to a single domain (a total test score). For most purposes, questions of medium difficulty are preferred, but question difficulty will always be a function of the nature of the content tested.

2) Correlations--One useful statistic for evaluating the adequacy of test questions is the question-test correlation. This index can indicate to the developer the extent to which any individual question is measuring the same thing as other questions in a cluster or in the total test.

3) Analysis of options--It is useful to determine the number of students who chose each option for a question and to note whether good or poor students were attracted to each option. This method of analysis can reveal ambiguities in the questions and possibly suggest other ways of testing the same concept.

Final Test Assembly: 1) Components of test assembly--Final test assembly encompasses activities such as the review, selection, revising, editing, formatting, and organizing of the questions or exercises for the final examinations.

2) Clarifying final responsibility--One individual should have primary responsibility for each examination regardless of the number of people contributing to the process and whether or not an outside group has contracted for the task. The involvement of a continuing committee is recommended at the time of final test assembly as it is at this point that all earlier work is synthesized. The final test assembler, though, needs to make the many decisions which will come up as the test nears completion.

3) Final item review--What precisely are the tasks facing the responsible individual, his cooperating staff members, and the committee? One significant task is final question review. All information available about the questions in the pool should be collected in a convenient form and each question reviewed in the light of this information. One useful strategy in this respect is the preparation of spiral notebooks with questions on one page and the following on facing pages:

- objective or area of specifications covered
- pretest information (if any)
- previous reviewers' comments
- correct response or scoring guide

Whenever questions are to be reviewed, it will be useful to keep the correct response separate from the text of the question so that the reviewer

can choose or formulate an answer and then check it against the official key or scoring guide. If the individual with primary responsibility for a test concludes that a question or exercise is ambiguous or that it lacks a single correct key, that question should not be used in a test, regardless of the quality of its pretest performance.

4) Meeting assembly specifications--The process of screening out questions because they are judged to be inadequate by reviewers can have the effect of reducing the pool of questions in some areas so that it appears impossible to meet the original specifications for a test. At this point, it is necessary to consider whether some previously rejected questions can be revised, whether additional materials can be created, or whether the intended scope of the test will have to be reduced. If it proves necessary to narrow the focus of a test, it will be important to describe just what is and is not being measured by the examination that is used operationally. In rare cases, the screening and culling process may produce a significantly larger body of questions than is needed to meet specifications. In such instances, one can sample from the pool in such a way as to leave a set of questions that is approximately equivalent to the questions used, thus creating the possibility of a parallel form for subsequent use.

5) Producing the final test--When organizing questions into a total test or set of related exercises, it will be useful to consider the impact of any actions on the nature of the final test. Decisions regarding the layout of questions on pages and their order and sequence may have considerable implications for the total cost of producing the final test. No test developer will be comfortable with page layouts that introduce complexities beyond those necessitated by the nature of the task. The use of type too small to permit easy reading or excessive packing of questions into pages may undercut the most careful effort to produce quality examinations. The individual with primary responsibility should review the final tests before they are reproduced. It is at this point that one is likely to discover that stimulus and response materials have been inadvertently separated, the options for multiple-choice questions are improperly sequenced, or that no space was provided for students to respond to free-response questions. (For a detailed discussion of the test production process, see Thorndike [7]).

6) Documentation--Although the task of test assembly is often so complex and demanding that it is difficult to set aside the time to keep accurate records of decisions made, the absence of such records can often create substantial problems for future test development. In general, every effort should be made to pick up potential errors as early as possible in the development process, so that last-minute changes without a significant number of later reviews can be avoided. There will always be a need, however, for changes to correct errors that are discovered at the eleventh hour. Careful documentation of the reason for the change, the nature of the change, and the steps that were taken to inform all significant people will reduce the probability of catastrophic errors. Imagine the consequences of rearranging the questions for a test at the last minute without notifying the individual who has already prepared the official scoring key!

Administration and Scoring: The same care that is given to developing a test should be devoted to planning for and holding test administrations. Responsibility for maintaining standard conditions, observing agreed-upon timing, and so on falls upon the assessment coordinator. In the case of essay examinations, scoring sessions must be held that permit reliable judgments of students' competencies (1).

If Our District Decides to Construct Examinations, How Much Use Should We Make of Essay Examinations?

The discussion of examination-development procedures earlier in this chapter included references to essay questions because such questions are widely used in schools. Essay tests are often employed by teachers who value the fact that such examinations require a student to develop answers from his own background, without the benefit of suggested possibilities, and to express the answer in his own words. Essay tests may also seem easy to prepare since a relatively small number of questions are used. Moreover, the use of essay questions very largely eliminates guessing.

It is true that essay questions can serve to measure some higher-level abilities when pupils are required to present evidence, evaluate, analyze, solve new problems, or approach problems in a new way.

Unfortunately, many essay questions do not do that. These are of the "Name the six largest cities in the United States" or "List the characters in each of the following stories" variety. Essay questions are often ambiguously stated or stated so generally that a pupil can bluff or "talk around" the subject. Poor essay questions of the type described are easy to write. Writing more challenging essays requires considerable thought.

A major disadvantage of the essay question is that the method of scoring the answer is so unreliable. Why is it so difficult to achieve reliability in scoring the answers to essay questions? One reason is that judgments differ. One grader may think that an answer is good; another may think it is poor. This may not be the result of the inadequacy of one of the graders; it may be the result of an honest difference of opinion. For another thing, if the essay question is concerned with a controversial topic, the grader's judgment is likely to be influenced by his own convictions. In the classroom setting, some teachers believe that the pupil's ability to express himself must be taken into account. These teachers deduct credit on answers to essay questions in social studies, science, and other subject matter areas for poor English expression. Other teachers believe that a pupil should not be penalized if he knows the subject matter but cannot express himself especially well. Then, too, the teacher's judgment may be influenced by the halo effect. There may be a tendency to grade in terms of work in class or even in terms of what is expected of the class to which the pupil is assigned. The halo effect may also operate from question to question, the quality of the answer to the first question influencing the scoring of the answers to subsequent questions. Add to all of these sources of unreliability of scoring the fact that teachers may correct tests after a long school day, even late in the evening, and that their scoring is likely to be somewhat erratic as a result of real fatigue. Sometimes teachers who rescore essay questions after an interval of time find themselves coming up with quite different scores.

Beyond the problem of reliable scoring, there are two other limitations to essay questions that should be mentioned. First, there is the likelihood of inadequate sampling of the subject when only a few questions are selected to cover a large content area. Then there is the large penalty per question that will result if the pupil does not know the answer to the question, or even more serious, if he knows the answer but does not understand the question.

In order to give students an opportunity to write in response to a question they are comfortable with, a choice of essay questions is frequently permitted. If the performance of pupils is to be compared one with another, a choice among essay questions should not be permitted. Without elaborate equating of the questions on the basis of their difficulty to the pupils who attempt them, there is no way of knowing, for example, how well the pupils who chose to answer questions 1, 2, and 4 would have done had they chosen to answer questions 3, 5, and 6 instead. And the better pupils who attempt the more challenging questions sometimes write less acceptable answers than do the less talented pupils who are satisfied to answer the easier questions.

One possible method of evaluating the adequacy of the essay scoring arrangements within a school would be for the School Evaluation Committee to arrange periodically for an independent scoring session. All the essays submitted over a period of time could be collected, read, and noted by a panel of judges. The results of this special session could be compared with those obtained at the departmental level.

In summary, essays are reliable assessment devices only when multiple, independent grading is arranged, using graders who can agree in advance on common criteria. For a comprehensive discussion of essay examinations, see Coffman (1).

What About Developing Local Objectively Scorable Examinations?

Questions of the objective type have at least four advantages over essay questions:

1. They permit wider sampling of learning in a relatively short time, making spotty preparation on the part of the pupil more obvious.

2. They can be reliably scored. If the items are unambiguous and the test has been keyed properly, the scoring errors will be clerical errors rather then errors of judgment.

3. They are more easily scored. Scoring time is reduced. The teacher is freed from suspicion of partiality.

4. They lend themselves rather readily to statistical analyses that can improve the quality of questions. The teacher can, over a period of time, assemble a file of questions, retaining from each test the questions that are of proper difficulty and that discriminate well between high and low achievers.

One limitation of questions of the objective type is that they are difficult to construct if they are to test anything more than memory. All types of objective questions, especially multiple-choice questions, can test for recognition

of assumptions, interpretation of data, recognition of limitations, application of principles, and a variety of other higher intellectual abilities and skills. Experience with several types of objective questions, however, suggests that some are more efficient for such purposes than others. Because the construction of objective questions that do measure understanding and thinking requires a large amount of time and considerable ingenuity, teachers are likely to be content with questions testing principally knowledge of facts--sometimes very trivial facts. This limitation of questions of the objective type, then, is not a limitation inherent in the questions themselves but a limitation imposed by the test constructor.

Five types of objective questions are popular with classroom teachers--completion, true-false, matching, classification, and multiple-choice. All of these types can serve useful purposes if the questions are well drawn and the answers to be supplied or chosen are definite and specific. They are all easily adapted to classroom situations, and a large number of them can be administered in a relatively short time. In most testing situations, a balanced blend of several types of objective questions will prove most effective. When more than one type is used, however, it is a good idea to group together questions of the same type so that the pupils will not be confronted with too-frequent changes in directions.

Final Thought

The foregoing discussion of essay and objective examinations was designed in part to illustrate the complexity and demanding nature of the job of developing good tests. The discussion contained in this chapter and the references to other material should provide Massachusetts educators with the tools to handle such a job. Many school and community people will need to make contributions over a period of time, but it is possible to achieve the goal of a cohesive and fair assessment system for nontraditional and experiential learning programs.

REFERENCES

1. Coffman, W.E. Essay examinations. In Robert L. Thorndike (Ed.) Educational Measurement (2nd Ed.) Washington, D.C.: American Council on Education, 1971.

2. Ebel, R.L. Essentials of educational measurement. Englewood Cliffs, N.J.: Prentice-Hall, 1972.

3. Finley, J. & Berdie, F.S. The national assessment approach to exercise development. Ann Arbor, Michigan: National Assessment of Educational Progress, 1970.

4. Kibler, R., Barker, L. & Miles, D. Behavioral objectives & instruction. Boston: Allyn & Bacon, 1972.

5. Mager, R. Preparing instructional objectives. Palo Alto, Calif: Fearon Publishers, 1962.

6. McAshan, H.H. Writing behavioral objectives: A new approach. New York: Harper & Row, 1970.

7. Thorndike, R.L. Reproducing the test. In Robert L. Thorndike (Ed.) Educational Measurement (2nd Ed.) Washington, D.C.: American Council on Education, 1971.

8. Tinkelman, S.N. Planning the objective test. In Robert L. Thorndike (Ed.) Educational Measurement (2nd Ed.) Washington, D.C.: American Council on Education, 1971.

9. Vargas, J.S. Writing worthwhile behavioral objectives. New York: Harper & Row, 1972.

10. Wesman, A.G. Writing the test item. In Robert L. Thorndike (Ed.) Educational Measurement (2nd Ed.) Washington, D.C.: American Council on Education, 1971.

CHAPTER 12

LOOKING BACK AND LOOKING AROUND

The previous chapters of this Handbook have outlined management procedures for designing and initiating programs for the evaluation of nontraditional and experiential learning at the secondary school level in Massachusetts. The authors have attempted to draw on the experiences of programs already in existence in the Commonwealth of Massachusetts and elsewhere in the country in order to raise possibilities and encourage interest in trying out new learning programs and new ways of evaluating the students who participate in them. We hope that many of the ideas that have been presented will be applicable not only to the nontraditional learning setting but to many aspects of the regular school program. We feel that a thoughtful and systematic approach to evaluation can make a considerable contribution to any educational program. To the extent that our attempts have been successful, the Handbook should contain material that will be of value even to those schools that choose to make minimal use of nontraditional learning approaches as part of their high school offerings.

Although we have included a good deal of material in this Handbook, we recognize that no single document can provide everything a school needs to develop an effective evaluation program. The references we have provided should help school staff members who wish to go beyond what is contained here. It should be noted, though, that no amount of studying of existing work will substitute for the experience that school districts' staff will acquire as they attempt to design and implement their own evaluation systems for nontraditional or experiential learning programs. Instead, school district staff need to prepare for an ongoing pattern of planning, try-out, review, and recycling. Wherever possible, programs should be initiated on a pilot basis involving only small numbers of students in certain learning areas. Any new educational endeavor in the schools will encounter complications and side effects not anticipated by even the most careful program planners. By starting in a small way and making explicit provisions for profiting from early experiences, school district staff can arrive at procedures appropriate to their school before investing large amounts of staff time and school resources in the activity. A school that tries to move too far or too fast in the direction of evaluating nontraditional learning programs will incur the risk of raising unrealistic expectations and promising more than can be delivered. By "making haste slowly," a solid foundation can be built for further expansion of a tested and accepted set of evaluation procedures for alternative learning programs.

The Need for Evaluation

Today seems to be a time for questioning the policies, procedures, and purposes of the secondary school. The Task Force on Secondary Schools in a Changing Society of the NASSP, for example, lists the following questions about secondary education in its 1975 report, This We Believe (1): What is its purpose? How should it function? Who are the appropriate participants? What constitutes a relevant curriculum and methodology? What revisions must be made?

Although the report of the Task Force and those from other recent conferences and commissions agree on the central goal of fostering student learning and development, they recognize few if any unalterable "givens" of secondary education. While some proposals are made to lower the school-leaving age, others are made to increase the length of formal schooling but to make out-of-school learning experiences the basic pattern for the final years of the secondary school program. Whereas some groups seek to place greater emphasis on minimum standards for school attendance in awarding the high school diploma, others recommend the shift to a competency-based approach that would make it possible for many students to earn the diploma, and thus complete formal secondary school education, by age 15.

We do not pretend to know which directions will be pursued in the future by most schools in the Commonwealth, nor can we anticipate whether productive resolutions will be found to such problems as providing all students with basic skills while giving students the opportunity to pursue diverse paths toward their own learning goals. We do feel, though, that whatever the course taken by a particular school program, or student, there will be a continuing need for soundly conceived and executed evaluation programs. It is our hope that this Handbook will be helpful to the Massachusetts educators who design and develop such programs and ultimately, to the students whose learning is evaluated.

REFERENCES

1. This we believe: Secondary schools in a changing society. Reston, Virginia: National Association of Secondary School Principals, 1975.

APPENDIX A

Sample Interview Schedule

Initial Needs Assessment Survey

Introduction

The following questions deal with some new programs and new approaches that the school is considering. I will be asking you first about your views on the general need for these programs. Then I will be asking you about your own (or your son or daughter's) particular interest in these programs.

General Needs (Teachers, Parents and Students)

Would you take a minute to look over this card. This is a list of some types of students who might need or benefit from some new and different programs. These new programs would take the student outside of the classroom and usually outside of the school.

Students	New Programs
1) artistically-oriented	—work or study in art museums, conservatories, or studios or with theatre groups
2) skill-oriented	—work or train in auto or TV repair shops, apprenticeship programs, vocational-technical centers
3) service-oriented	—volunteer service in government, hospitals, agencies, etc.

Students	New Programs
4) science-oriented	—work or study in libraries, museums, zoos, laboratories
5) restless	—some different learning arrangements
6) interested in learning better work habits	—work-study-counseling programs focusing on work habits (dependability, punctuality, etc.)
7) career explorers	—brief work study experiences in varied types of work
8) career beginners	—work-study in career path (e.g., father's business, beginning secretaries, etc.)
9) independent learners	—special study projects outside of class assignments
10) special opportunities	—program based on travel or other opportunities

—Are there any other kinds of students, not shown on this list, who would benefit from some new program?

—Based on your knowledge of the students here, which of the above types of students would include the greatest number? the smallest number? How many students would you estimate would be in each of these categories? (Include your own added categories.)

—In your opinion, are all of these programs in keeping with the real purposes of the school? Which of these programs should the school consider first?

—If these new programs are not introduced (complications, budget problems—a lot of things can happen!) would any of these types of students suffer some real hardship? If so, please describe the hardship.

Personal Needs (Students and Parents only)

—Did you see yourself* in any of the groups listed above (including your own additions to the list)? (*for parent, "your son or daughter")

—If several categories are cited, please rank these categories in order of their importance to you! The following questions will refer to the category which you rated most important. If no categories are cited, proceed to final question.

—Describe the basis for your need for this program (e.g., previous experience, special interests, etc.)

—What do you feel you would miss if you did not take part in this program?

—Describe how you would imagine that this program would work?

—What do you see to be the advantages? disadvantages of such a program?

—How would this new program be linked to what you are doing in school now?

—What, if any, kinds of requirements should the school set for this new program? What about the evaluation of outcomes (reports, products, tests)? Supervision? Selection of those eligible to take part? Counseling?

—In closing, could you please tell me a few things about yourself? Your name? Your age? What year are you in school? Types of program you are taking? Present plans for after high school? Subjects you like best? least? Your general rating of your high school experience?

—Do you have any other comments or ideas you want to add to this discussion? If not, thank you.

APPENDIX B

New York External High School Diploma Program

The External High School Diploma Program, guided in its design by a set of assumptions about adult learners, provides the opportunity for adults to demonstrate real-life competencies in practical settings. The Regional Learning Service of Central New York provides the assessment staff and the "learning consultants."

The first step in the program is a diagnostic pretesting administered by a learning consultant, which helps the learner identify competencies already possessed and those in which he or she is deficient. If pretesting reveals that competencies are at the required level, the learner begins the second step: the assessment of Generalized Competencies. A task force identified and by consensus selected 64 competencies which responded to the question: "What skills do adults need to cope with daily living?" The content of the assessment procedures for Generalized Competencies consists of projects based on the practical developmental tasks of adulthood (changing residence, applying for a job, consumer awareness, caring for family and health). These projects are performed as take-home simulations and oral interviews, thus providing flexibility in time and place.

In addition to demonstrating a common set of minimum general competencies, each candidate must also prove possession of advanced competencies in an area of his or her own choice. Thus, the third step in the process is the assessment of Individualized Competencies in either occupational/vocational skill, advanced academic ability, or advanced specialized competency such as drama or art.

The candidate can choose his or her own assessment mode from alternative assessment procedures used to evaluate advanced competencies. In this way, the program introduces an element of personal freedom and control. Occupational/vocational competence is assessed by either (a) employer testimony, (b) performance as judged by a specialist, or (c) possession of a New York State license in a particular area. Advanced academic competence is assessed by either (a) original work indicating college-level ability, or (b) an acceptable score on a standardized test measuring readiness for college-level work. The advanced specialized competence assessment procedure, designed primarily to enable persons to demonstrate a strong vocational skill, is handled on the basis of judgment by specialists.

Once a person has demonstrated possession of Generalized Competencies as well as Individualized Competence, the Regional Learning Service assessment team submits the student's portfolio for review by the Regional Committee composed of 16 local persons from business, education, and industry. Upon their recommendation, the Cortland New York School Board awards the Regents External High School Diploma.

Within one year after program development, more than 25 adults ranging in age from 19 to 61 have met the requirements for the diploma. A follow-up study of the recent graduates revealed that the possession of the diploma had a number of salutory effects. For example, a number of students were able to take the Civil Service examinations, a prerequisite of which is the high school diploma; three students entered a community college; and four persons were able to retain their jobs which had been in jeopardy due to a newly imposed high school diploma requirement.

Further information regarding the external diploma may be obtained by writing to the Project Director, Central New York External High School Diploma Program, 405 Oak Street, Syracuse, New York 13203.

APPENDIX C

Dynamy High School Internship Program

(Worcester Public High Schools)

Dynamy is an educational organization based on the ideal that significant learning and growth may occur through direct, personal experience in settings outside the classroom. The basic component of the Dynamy High School is the internship which provides a structured working and learning experience, paid or unpaid, in the community. Students enter Dynamy in lieu of their senior year and begin their Dynamy experience with a three-week outdoor leadership experience. They continue in an internship consistent with their needs, abilities, and interests. An admission committee composed of parents, Dynamy staff, and school personnel selects candidates on the basis of motivation and evidence of the maturity needed for performance in the Dynamy year.

The form and content of the internships vary widely. Some are typical apprenticeships involving the acquisition of particular skills in a craft or trade; others begin as an exposure/observation experience which may lead into a project activity in an area appropriate to the student's skills and interest. All the internships share the following elements:

1. A sponsor who is the on-the-job supervisor to whom the student is accountable

2. A contract, mutually agreed upon by student and sponsor, which specifies the goals, expectations, tasks, and responsibilities of the student

3. A task performed by the student, important to the company, and not simply generated to keep the student busy

4. A written evaluation by both sponsor and intern at the conclusion of the experience

The goals of the Dynamy program are to (a) develop the ability to work well with people, (b) help young people achieve independence and assume responsibility for their own educational and personal development, (c) provide the opportunity for high school students to "try on" different career and occupational roles, and (d) to enhance overall decision-making ability of participating students.

Positive sponsor evaluation and the extent to which the student meets the goals specified by his or her contract provide the basis for the award of academic credit.

Inquiries regarding the Dynamy High School Program may be addressed to: Dynamy, Inc., 850-852 Main Street, Worcester, Massachusetts 01610.

APPENDIX D

Executive High School Internship Program

The Executive High School Internship Program enrolls over 2,000 students in 28 cities and towns across the country. The interns are juniors or seniors who take a semester's or quarter's leave of absence from regular studies and serve, without pay, as special assistants to senior officials in organizations in practically any field in which an appropriate internship may be developed. The young executives attend policy meetings and conferences with their sponsors, follow up on special assignments, prepare memoranda and reports, and, in a few instances, travel with their sponsors. Interns observe and interact with their sponsors four days a week, and on Fridays meet with the high school coordinator for seminars on urban policy and development, some of which the students plan themselves.

Through the program, students are able to explore a variety of career clusters and gain specialized knowledge in the field in which they are placed. Furthermore, through peer relationship with adults, executive interns may experience growth in self-confidence as well as improvement in the communications and decision-making skills.

The students receive academic credit for their experience and must keep daily analytical logs of their activities. Also, at the end of the semester, they are required to present a project to their high schools demonstrating what they have learned that semester.

The Executive High School Internships of America, underwritten by several foundation grants, provides a full range of staff training and technical assistance services to each subscribing school district.

Inquiries regarding the program may be addressed to Executive High School Internships of America, 680 Fifth Avenue, Ninth Floor, New York, New York 10019.

APPENDIX E

The Junior Achievement Program

The Junior Achievement Program, begun in 1919 at Springfield, Massachusetts, was developed in recognition of the need for increased economic awareness on the part of high school students. Through the program, thousands of young people are learning how to manage minature businesses under the guidance of adult advisers drawn from the business and service industries. Junior achievers and their advisers meet two hours, one night a week for 32 weeks at their area Junior Achievement Center. Based upon the educational nature of these meetings, the students form a company, sell stock to relatives and friends, and elect company officers; then, under the guidance of the adult adviser, they manufacture and sell their product or service. At the end of the school year, the students liquidate their business and issue an annual report of profit or loss to the stockholders.

Through hands-on economic and business experience, junior achievers gain practical business skills, working knowledge of economics, experience in decision making, leadership, and human relations. The extent to which these objectives are met can be assessed in a number of ways, the most formal of which is a set of examinations which one may take to demonstrate what has been learned about business and economics in general, and about the particular role one has performed during the Junior Achievement year. Nineteen schools in Massachusetts are now granting credit for satisfactory Junior Achievement experience.

Recently, Junior Achievement of Eastern Massachusetts has drafted a proposal which offers high schools the opportunity to incorporate the Junior Achievement Program, as outlined above, into the existing curriculum and to supplement existing courses with practical experiences in economics. Implementation of the proposal would incorporate the following elements:

1. The school will establish Junior Achievement companies within selected courses that are presently offered in the curriculum.

2. The Junior Achievement Company will meet for two class periods each week for 30 weeks.

3. In addition to the adult adviser provided by the sponsor company, the classroom instructor will act as an adviser to the Junior Achievement Company.

4. Junior Achievement will provide the necessary materials and advice needed to assure smooth operation of the program, including (a) training of the teacher-coordinators and "company" officials,

(b) provision of books, records, and kits which give the company a week-by-week agenda, and
(c) supervision and planning of contests, awards, scholarships, and conferences.

5. Junior Achievement will also contact and obtain the services of counseling firms (local businesses).

Junior Achievement is a practical business-economic educational program which demonstrates for young people the mechanics, the benefits and the value of a free enterprise system. Understanding the economic system under which American business operates is an important asset, for everyone must participate in the system as a manufacturer, a business executive, or, at least, as a consumer.

Inquiries regarding the Junior Achievement Program may be addressed to the Junior Achievement of Eastern Massachusetts, Inc., 256 Huntington Avenue, Boston, Massachusetts 02115.

APPENDIX F

The Home Base School

Watertown, Massachusetts

The Home Base School (HBS), a public alternative high school in Watertown, Massachusetts, opened its doors in September 1971 with 100 students and seven staff members. Based on the assumption that students can learn in many places outside school buildings, the scope of education at Home Base varies from basic courses in mathematics, science, language arts, and social studies, to unique experiences in the community including experimental theatre, music theory, photography, journalism, and karate. More than 100 resource people--businesspersons, parents, college students and professors--teach, give tutorials, or provide apprenticeships to HBS students.

A unique feature of HBS is that the students, in collaboration with parents and the staff, construct the program offerings of the school. This policy is consistent with the HBS ideal that since students must live with the consequences of decisions, they should be allowed to play a significant role in making them. In view of the emphasis on student-initiated programs, an adviser system is the core of the school's structure. Through the adviser system, whereby each student meets weekly with a particular staff member, students receive individualized guidance in their movement towards taking responsibility for their own education, initiating independent action, and developing skillful decision-making ability.

Throughout their experience at Home Base, students have the opportunity to investigate careers and to receive credit for work done in the community as long as they have the experience monitored by a staff member. The process of evaluating the outcomes of these as well as the more traditional experiences is three-fold:

1. Monthly memos are written by the teacher/supervisor of the course or project to the adviser. The memo provides an appraisal of the students' work and attendance and encourages on-going communication among adviser, teacher or supervisor, and student.

2. A final evaluation is written by the teacher or the supervisor which includes the requirements and expectations of the course or work experience and an assessment of how well the student accomplished the requirements and met the expectations. The final evaluation includes a grade of pass or fail as well as comments prepared in collaboration with the student concerning his or her strengths and/or weaknesses, progress and development, participation and attendance, effort, and consistency.

3. The transcript is a summary of the course or project evaluations completed throughout the year. The transcript statement, compiled by the students' adviser, provides a thorough record of the specific work that the student has accomplished.

Home Base School conducts an intensive recruitment program in order to provide a sufficient number of applications for a stratified random-selection procedure. The random choice ensures fairness in the selection process and the stratification enables Home Base to maintain a cross-section of abilities and interests among its student body that reflects the other Watertown secondary school populations. The applications from which the selection is made are those which evidence positive interest in alternative high school education and which indicate a reasonable probability of succeeding at Home Base.

Further information regarding Home Base School may be obtained by writing the Program Chairman, Home Base School, 465 Mt. Auburn Street, Watertown, Massachusetts 02172.

APPENDIX G

Metropolitan Cultural Alliance of Greater Boston

Since its founding in 1968, the Metropolitan Cultural Alliance of Greater Boston has grown to include over 100 organizations such as museums, zoos, historical societies, orchestra, opera, dance and theater companies in addition to many of the educational institutions offering programs in the arts.

In 1975 the Alliance's Educational Project involved 1,300 students at all grade levels in the metropolitan schools. A series of 14 pilot projects designed to facilitate racial integration and funded by the State were conducted by the Alliance in preparation for expanded cultural programs in the schools in 1976.

One of these pilot projects brought 80 high school students together from five schools. In a history and humanities course entitled "The American Experience," high school teachers of English and social studies coordinated classroom study with outside-of-classroom activities. Teachers accompanied groups of 20 students, half from inner-city and half from suburban schools, once a week for six weeks to a full-day program at the Museum of Fine Arts and the Museum of Transportation. They observed specific exhibits related to the day's topic such as vehicles, furnishings, utensils, costumes, documents or art objects, and saw slide presentations. Live presentations were also given by four performing arts companies. Artists from the Concert Dance Company, the Pocket Mime Theater, Stage I, and the Theatre Workshop of Boston instructed students in the techniques of dramatization and role playing to help them learn about the course content through a different medium. Students presented individual and group projects as well as preparing academically for the day's subject. On the basis of the fulfillment of the general goals of the programs which were planned during advanced teacher workshops, students were given credit for their participation. Assessment and evaluation were done by the teachers and school program coordinators, and the schools generally applied credit on an individual basis. The objectives were not only to advance students' comprehension of the history and culture of people of another era to whom they could relate many of their own experiences, but also to stress the importance of mutual respect for one another as young people from different racial and economic backgrounds. The observing students learned together on "neutral territory" (in the museums and art galleries), which the majority of urban and suburban students alike had never visited. Teachers were asked to evaluate student attitudinal changes, and in this connection to assess student's reactions to a program which exposed them to many new ways of utilizing the cultural resources of the city and which produced more positive attitudes towards the arts than are often present at high school age. In addition, schools could give valid academic credit for this in-depth learning experience.

The most important institutional objective was to build the foundation for expanding the number of combined history and humanities courses that make use of cultural organizations throughout the State as part of the regular school curriculum.

The Metropolitan Cultural Alliance has published a booklet called Culture Cracks the Blackboard, which lists most of the resources of the Boston area. The Alliance also publishes a monthly newsletter called Currents. To contact the Alliance, write to 37 Newbury Street, Boston, Massachusetts 02116.

APPENDIX H

Sample Documents for a Performance Test*

Part I Instructions to Observers

Part II Instructions to Students

Part III Observer's Rating Form

*Sample documents taken from chapter 5 of Boyd, J. L. Jr. & Shimberg, B. Handbook of performance testing, a practical guide for testmakers. Princeton, New Jersey: Educational Testing Service, 1971. Used with permission.

INSTRUCTIONS TO OBSERVERS FOR A PERFORMANCE TEST

THE START CIRCUIT OF A 5-KW GENERATOR SET (PART I)

1. The administrative details concerning this test have been prepared in three parts. This is Part I. Part III is the rating form to be completed by the observer during the student's performance. It includes a listing of all the items to be judged. Part II is instructions to students. Observers are not to assist the students during the Performance Evaluation. They are to keep close surveillance of the student's actions for evaluation purposes and also to prevent accidents or damage to the equipment. Observers will evaluate the student on performance of requirements shown below. If the student successfully completes a step of a requirement, he is to be awarded the total amount of points for that step. <u>If he does not successfully complete a step, he is to be given a zero for that step</u>.

   a. This test is to be administered indoors, under artificial light.

   b. Cables and load banks will be connected for 120 volts, single phase, 60 cycle operation prior to the start of the Performance Test.

   c. Wire #P29B18 going from #2 terminal of the S6 switch to terminal #2 of the ammeter will be removed and replaced by a false wire.

   d. The following equipment will be furnished the student at the site:

      (1) 5 KW generator
      (2) Multimeter
      (3) Hydrometer
      (4) Tool kit
      (5) Schematic
      (6) Load bank
      (7) Connecting cables
      (8) Grease pencil

e. When the student completes each detail noted on the performance rating form, record his score in the appropriate column. He will be graded on the following:

(1) Attempt to start engine electrically
(2) Check battery voltage and specific gravity
(3) Disconnect battery
(4) Trace engine start circuit on schematic diagram using grease pencil. The student will take each part of circuitry in sequence: (a) energize the K4 coil,(b) energize the K1 coil, (c) energize the K4 coil and (d) energize the start motor.
(5) Make continuity checks of the start circuit with a multimeter
(6) Point out the defect to the instructor on both the equipment and schematic

f. The maximum time allowed for the test is 30 minutes.

INSTRUCTIONS TO STUDENTS FOR A PERFORMANCE TEST
THE START CIRCUIT OF A 5 KW GENERATOR SET (PART II)

1. The purpose of the test is to provide you with an opportunity to demonstrate your ability in troubleshooting the start circuit of the 5 KW generator. Using the prescribed procedure you have been taught, you are to find any incomplete circuitry.

2. You will be allowed 30 minutes in which to complete the test. A satisfactory grade cannot be achieved unless the correct procedure is followed.

3. During a field exercise you are required to furnish power for lighting the area. The output power is 120 volts, single phase, 60 cycle alternating current. You have already performed the required Preventive Maintenance Services and you found no deficiencies.

4. The following equipment will be provided:

5 KW generator
Load bank
Load cables
Multimeter
Hydrometer
Tool kit
Schematic
Grease pencil

5. Requirements:

a. Attempt to start engine electrically (10 points)
b. Check battery voltage with multimeter and specific gravity with hydrometer and orally report to instructor (16 points)
c. Disconnect battery (5 points)
d. Trace the engine start circuit on the schematic diagram using a grease pencil, starting at (-) side of battery (22 points)

e. Make a complete continuity check of the first part of the start circuit to make sure that K4 coil can become energized, with a multimeter (47 points)

f. Point out the defect to the instructor on both the equipment and schematic (20 points)

6. On completion of troubleshooting the circuit, move to the next station.

OBSERVER'S RATING FORM FOR A PERFORMANCE TEST

THE START CIRCUIT OF A 5-KW GENERATOR SET (PART III)

Student_____Time Started_____Time Completed_____

Possible Score 120 Critical Score 84 Student Score_____

	Score if Task Properly Accomplished	Student Score
1. ATTEMPT TO START ENGINE ELECTRICALLY		
a. Place fuel valve in tank position	1	
b. Place oil pan baffle rod in proper position	1	
c. Place air cleaner intake shutter in proper position	1	
d. Place governor control in start position	2	
e. Place remote/local switch in local position	1	
f. Place emergency run/stop switch in normal position	1	
g. Hold start switch in start position (no more than 15 second intervals)	3	
TOTALS	10	

This Observer's Rating Form covers five other areas:

2. Check battery voltage and specific gravity
3. Disconnect battery (negative ground)
4. Trace engine start circuit on schematic
5. Make continuity checks of start circuit with multimeter
6. Point out the defect to the instructor on both the equipment and the schematic

SELECTED MACE STUDIES

Title	Description	Available From
Report of the Massachusetts Business Task Force For School Management (Warren King and Associates)	105 ways to save money in school management	Submitted to ERIC
Approaches to Collaboration: A Handbook of Strategies for Serving Special Needs Children (Education Collaborative for Greater Boston)	Processes for planning special education collaboratives	Education Collaborative for Greater Boston 186 Hampshire St. Cambridge, MA 02138
Effectiveness, Efficiency & Equal Opportunity in the Public Schools of Massachusetts (Governor's Commission on School District Organization and Collaboration - Summary)	Specific recommendations on regional evaluation, sharing resources, and organizing districts	MACE ERIC (ED 097-755)
The Role of the Superintendent of Schools in Massachusetts: A Handbook for School Committees and Superintendents (FITZGERALD)	Survey of laws and factors affecting productivity of superintendents	Massachusetts Association of School Superintendents 73 Tremont St.,Rm.1114 Boston, MA 02108 ($1.50)
Organizing an Urban School System for Diversity (CRONIN)	Recommendations for organization of the Boston public schools	D.C.Heath Publishing Co. Lexington, MA 02173
Higher Education in Massachusetts: A New Look at Some Major Policy Issues (Academy for Educational Development - TICKTON)	Six major recommendations on policy in Massachusetts	Summary - MACE Full - ERIC (ED 083-905)
Strengthening the Alternative Post-Secondary Education System: Continuing and Part-Time Study in Massachusetts (University Consultants) (NOLFI & NELSON)	Vol.I makes recommendations on access to educational services Vol.II inventories programs and program relationships	Vol.I - ERIC (ED 002-816) Vol.II- ERIC (ED 002-817)
The People's Colleges: The State Colleges of Massachusetts (COLLINS et al)	Recommendations for the development of liberal arts and career programs other than teacher education in 11 state colleges	ERIC (ED 065-069)
Guidelines for Planning and Constructing Community Colleges (DUNSMORE)	Relates educational criteria to guidelines on administrative personnel, space, curriculum development, and budgeting.	ERIC (ED 034-390)

SELECTED MACE STUDIES (continued)

	Title	Description	Available From
financing in education	Massachusetts Taxes: A Factual Guide to Future Action (Massachusetts Taxpayers Foundation -- ZIEGLER)	Presents data which can be used to predict revenue and taxation relationships.	MACE - single copies
	Financing Public Education in Massachusetts: A Process for Revision (HEFFLEY)	A statewide survey of opinions related to issues of finance reform	MACE
	Financing Public Schools in Massachusetts: Problems of the Past, the Present and the Future (NESDEC)	Proceedings of a conference sponsored by the Massachusetts Educational Conference Board	ERIC (ED 093-075)
	Aid to Private Higher Education in Massachusetts Why? How? (TERMAN)	A monograph submitted to the Academy for Educational Development for use in its report HIGHER EDUCATION IN MASSACHUSETTS: A NEW LOOK AT SOME MAJOR POLICY ISSUES	MACE ERIC (ED 081-333)
	The State Dollar and The Schools (RYAN)	The formulas which determine state aid to education and proposals for change	ERIC (ED 101-474)
	The Measurement of Alternative Costs of Educating Catholic Children in Public Schools (DANIERE & MADAUS)	A report which recommends partial state funding for parochial schools	ERIC (ED 043-943)
curriculum	Something of Value: A Summary of the Findings and Recommendations for Improving Elementary Science in Massachusetts (Office of Instructional Research and Evaluation, Harvard University -- WHITLA & PINCK)	A summary of a state-wide evaluation which found that NSF curricula have several compelling advantages over traditional, textbook approaches to teaching elementary science	MACE ERIC (ED 091-160)
	Lively Elementary Science: A Handbook of Suggestions for Introducing and Maintaining Innovative Science Activities. (Harvard University, Office of Research and Evaluation -- WHITLA & PINCK)	Translates the results of evaluative research into suggestions for improving elementary science programs through collaboration, in-service training and leadership development.	MACE
	Success and Failure: A Summary of Findings and Recommendations for Improving Elementary Reading in Massachusetts City Schools (Educational Research Corporation -- ELLIS)	An attempt to isolate school factors which might account for reading performance in urban elementary schools. While not identifying any simple answers, provides a basis for important, self-analysis in any elementary school.	MACE ERIC (ED 097-386)

SELECTED MACE STUDIES (continued)

Title	Description	Available From
The Here, Now and Tomorrow of Cable Television in Education...A Planning Guide (Creative Studies -- LEVINE)	A comprehensive how-to-do-it guide for use local districts	ERIC (ED 086-175) Creative Studies, Inc. 167 Corey Rd. Boston, MA 02146 ($2.00)
Pupil Services for Massachusetts Schools (LIDDLE & KROLL)	Recommendations for equalizing pupil services in Massachusetts schools	ERIC (ED 037-767)
Occupational Education for Massachusetts (SCHAEFER & KAUFMAN)	Recommendations for replacing "general" curriculum programs with more flexible vocational	Summary - MACE
Quality Education for the High Schools in Massachusetts: A Study of the Comprehensive High School in Massachusetts (MICHAEL)	An appraisal which finds that the comprehensive high school can meet the needs of all youth in a community most effectively	Submitted to ERIC
Massachusetts Study of Educational Opportunities for Handicapped and Disadvantaged Children (BLATT & GARFUNKEL)	Stresses the need for a master plan for meeting special education needs. This report was instrumental in the design of Chapter 766.	Summary - MACE Full - ERIC (ED 061-678)
Guide for Students and Parents on the Massachusetts Regulations Pertaining to Student Records (ALLEN)	A 7-page guide which can be adapted for notifying students and parents of their rights under the student record law	MACE Send a stamped, self-addressed envelope with your request.